Learning based energy management in multi-cell interference networks

Zhang, Xinruo

Awarding institution:
King's College London

LEARNING BASED ENERGY MANAGEMENT IN MULTI-CELL INTERFERENCE NETWORKS

XINRUO ZHANG

KING'S COLLEGE LONDON

2018

LEARNING BASED ENERGY MANAGEMENT IN MULTI-CELL INTERFERENCE NETWORKS

XINRUO ZHANG

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
AT
CENTER FOR TELECOMMUNICATIONS RESEARCH
DEPARTMENT OF INFORMATICS
KING'S COLLEGE LONDON

2018

Acknowledgements

This thesis would not have been possible without the assistance and guidance of several individuals. It is a pleasure to take this opportunity to express my sincere gratitude to all who in one way or another contributed to the completion of this study.

First and foremost, I would like to express my utmost gratitude to my primal supervisor, Dr. Mohammad Reza Nakhai, for his continuous support and insightful guidance throughout my Ph.D studies. His meticulous attitude as well as enthusiasm and devotion for research have an enormous influence on me. Without his innovative perspective on research directions or immense knowledge on learning and optimization for wireless communications, this thesis would not have been completed. I could not have imagined having a better supervisor for my Ph.D studies and I look forward to having more opportunities to cooperate with him in the future.

I would also like to express my deep appreciation to all my friends across the world and my colleagues at Center of Telecommunications Research, King's College London, who supported me in various ways for the past four years, especially during times of hardship. With their company and encouragement, these four years have become a precious, rewarding and unforgettable experience for me. In addition, I would like to sincerely acknowledge all members of staff in Department of informatics including Prof. Mischa Dohler, Prof. Abdol Hamid Aghvami and Prof. Arumugam Nallanathan for their inspiring lectures and valuable advice.

Last but not least, I would like to dedicate this thesis to my beloved parents, Yidu Zhang and Ning Zhang, for their eternal love, boundless patience and selfless support throughout these years. They are my role models and they have shaped me into the person I am today. Their love is always my motive force and this work would not have been possible without their support.

Table of Contents

Abstract . . . . . **iv**

List of Tables . . . . . **vi**

List of Figures . . . . . **vii**

List of Abbreviations . . . . . **ix**

List of Notations . . . . . **xi**

List of Nomenclature . . . . . **xiii**

Chapter 1 Introduction **1**
1.1 Motivation 1
1.2 Contributions of the Thesis 7
1.3 Outline of the Thesis 9

Chapter 2 Background **10**
2.1 Introduction 10
2.2 Mathematical Preliminaries 10
2.2.1 Convex Optimization 10
2.2.2 Lagrangian Duality and Karush-Kuhn-Tucker Condition . . . 14
2.2.3 Semidefinite Programming 15
2.2.4 Multi-cell Multi-user Downlink Beamforming 16
2.3 Multi-cell Interference Network and Cooperative Transmission 18
2.3.1 Channel State Information at the Transmitters 19
2.3.2 Coordinated Transmission 21
2.3.3 Benchmark Cooperative Beamforming Designs 24
2.4 Energy Trading and Smart Grid 26
2.5 Reinforcement Learning and Multi-armed Bandit Problem 27
2.5.1 Q-learning for Markov Decision Process 29
2.5.2 Multi-armed Bandit Problem 29
2.5.3 Variations of MAB problem 31
2.6 Related Works 32
2.7 Concluding Remarks 36

Chapter 3 Robust Outage Probability based Distributed Beamforming for Multicell Interference Networks 37
3.1 Introduction 37
3.1.1 Main Contributions 37
3.1.2 Organization 38
3.2 Robust Transmission in Multicell Networks with Probabilistic Constraints involving Statistical CSI Uncertainties 39
3.2.1 System Model and Problem Formulation 39
3.2.2 Optimization of Problem in (3.5) 41
3.2.3 Distributed Optimization of Problem in (3.12) 44
3.2.4 Fronthaul Signalling Overhead and Computational Complexity Analysis 49
3.3 Robust Transmission in Multicell Networks with Probabilistic Constraints involving Instantaneous CSI Uncertainties 51
3.3.1 System Model and Problem Formulation 51
3.3.2 Optimization of Problem in (3.31) 53
3.3.3 Distributed Optimization of problem in (3.49) 58
3.4 Simulation Results 61
3.5 Concluding Remarks 68

Chapter 4 An UCB Algorithm for Worst-Case Distributed Robust Transmission in Multicell Networks 69
4.1 Introduction 69
4.1.1 Main Contributions 69
4.1.2 Organization 70
4.2 System Model and Problem Formulation 71
4.3 Distributed Optimization of Problem (4.4) 73
4.3.1 Distributed Optimization of (4.5) for a Fixed c_i 74
4.3.2 UCB Algorithm for Finding the Globally Optimal c_i 79
4.3.3 Fronthaul Signaling Overhead and Computational Complexity Analysis 83
4.4 Simulation Results 83
4.5 Concluding Remarks 88

Chapter 5 A Bandit Approach to Price-Aware Energy Management in Cellular Networks 89
5.1 Introduction 89
5.1.1 Main Contribution 89
5.1.2 Organization 90
5.2 System Model 90
5.2.1 Energy Management Model 92
5.2.2 Downlink Transmission Model 93
5.3 Price-aware Energy Management 94
5.3.1 Problem Formulation 95
5.3.2 Reweighted ℓ_1-norm and Semidefinite Programming 95

5.4 Proactive Energy Management 97
5.5 Simulation Results 102
5.6 Concluding Remarks 108

Chapter 6 Adaptive Energy Storage Management in Green Wireless Networks 109
6.1 Introduction 109
6.1.1 Main Contribution 109
6.1.2 Organization 110
6.2 System Model 110
6.3 Adaptive Storage Management Strategy 113
6.3.1 Problem Formulation 113
6.3.2 SDP Optimization 115
6.3.3 Proposed Online Learning Algorithm 115
6.4 Simulation Results 119
6.5 Concluding Remarks 121

Chapter 7 Conclusions and future work 122
7.1 Thesis Summary 122
7.2 Future Research Directions 125

Appendix A Proof of Lemma 3.2.1 127

Appendix B Proof of Lemma 4.3.1 128

Appendix C Proof of Lemma 5.3.1 132

References 135

List of Publications 145

Abstract

The ever-increasing energy requirement incurred by future dense wireless communication networks has always been a challenging issue. Eliminating the inter-cell interference (ICI) is considered as a key factor for green communication whilst adapting to energy demand variations contributes to the stable cost-efficient operation of the system. This thesis focuses on learning-based energy management and interference control among base stations (BSs) using convex optimization methods in multi-cell networks.

The robust distributed coordinated approaches are proposed to solve aggregate transmit power minimization problem constrained by certain signal-to-interference-plus-noise-ratio (SINR) outage probabilities in the presence of imperfect channel state information. The intractable problem is first converted to a tractable form and then decomposed into independent sub-problems to be solved at individual BSs. The individual BSs gradually learn the ICI imposed from other BSs via sub-gradient iterations with a light inter-BS communication overhead.

Then, the problem of maximizing the weighted SINR requirements is investigated. The original problem is first converted into an equivalent total transmit power minimization problem for a fixed scale of SINR targets. Then, an upper confidence bound based algorithm is proposed to optimally and distributively scale the SINR targets based on per-BS power budget and coordinate ICI among BSs.

Next, a combinatorial multi-armed bandit (CMAB) inspired online learning algorithm is introduced to minimize the time-averaged energy cost at BSs, powered by various energy markets and local renewable energy sources. The algorithm sustains traffic demands by enabling sparse beamforming to schedule dynamic

user-to-BS allocation and proactive energy provisioning at BSs to make ahead-of-time price-aware energy management decisions.

Finally, in order to address the dynamic statistics of renewable energy supply, an adaptive strategy inspired by CMAB model for energy storage management and cost-aware coordinated load control is proposed. The proposed strategy makes online foresighted energy storage decisions to minimize the average energy cost over long time horizon.

List of Tables

2.1 CSI acquisition for different scenarios 20

3.1 Simulation parameters [1, 2] 62

4.1 Simulation parameters [1–3] 84

5.1 Percentage of exploration using smart scheduling 99
5.2 Simulation parameters [4–6] 104

6.1 Simulation parameters [4–7] 118

List of Figures

1.1 5G technical improvement over 4G. 2
1.2 Energy efficient 5G solutions. 3
1.3 An example of 5G dense heterogeneous network. 5
1.4 Scope of research on energy management in this thesis. 7

2.1 Some simple convex and non-convex sets. 11
2.2 Example of a convex function [8]. 12
2.3 A typical example of multi-cell multiuser interference network. 19
2.4 Illustration of levels of collaboration amongst BSs. 22
2.5 A typical smart grid system. 26
2.6 A typical frame of a reinforcement learning scenario. 28

3.1 Illustration of system scenario. 39
3.2 Flowchart of Algorithm 3.2.1. 49
3.3 An example of user distribution in a 3-cell network. 62
3.4 Comparison of total transmit power versus various SINR outage probabilities and error variances. 64
3.5 Comparison of total transmit power with $\rho = 0.3$ for the proposed strategy and a) outage probability based design in [9], b) ADMM approach in [10]. 65
3.6 Power variation of Algorithm 3.2.1 at $\gamma = 10$ dB target SINR for $M = 6, 8$ antenna elements per BS. 66

4.1 Flowchart diagram of the proposed UCB algorithm 81
4.2 Comparison of total transmit power for different designs. 85
4.3 Histograms of average SINR satisfaction ratio at $\gamma = 10$ dB of: a) non-robust power minimization design in [11], b) proposed robust strategy. 87

5.1 Illustration of downlink partial cooperation among BSs. 91
5.2 Flowchart diagram of proposed online learning algorithm 100
5.3 An exploration-exploitation trade-off model of smart scheduling . . . 101
5.4 An example of multi-user downlink simulation topology. 103
5.5 Normalized total energy cost of proposed strategy versus other designs at individual time slots at $\gamma = 15$ dB 105

5.6 Normalized total energy cost of proposed strategy without smart scheduling at individual time slots at $\gamma = 15$ dB 106

6.1 Illustration of downlink partial cooperation among storage-deployed BSs. The information flow is denoted by dashed lines and the energy flow is denoted by solid lines. 111
6.2 Illustration of proposed energy storage management strategy 114
6.3 Normalized total energy cost of the proposed strategy versus design in [12] at $\gamma = 15$ dB at individual time slots 119
6.4 Normalized total energy cost of proposed strategy at $\gamma = 10$ dB and $\gamma = 20$ dB at individual time slots 120

List of Abbreviations

3G	The Third Generation
3GPP	The Third Generation Partnership Project
4G	The Fourth Generation
5G	The Fifth Generation
ADMM	Alternating Direction Method of Multipliers
AoD	Angle of Departure
AWGN	Additive White Gaussian Noise
BS	Base Station
C-RAN	Cloud Radio Access Network
CAPEX	Capital Expenditure
CB	Cooperative Beamforming
CDF	Cumulative Distribution Function
CMAB	Combinatorial Multi-armed Bandit
CO2	Carbon Dioxide
CoMP	Coordinated Multipoint
CP	Central Processor
CS	Coordinated Scheduling
CSI	Channel State Information
CSIT	Channel State Information at the Transmitter
CUCB	Combinatorial Upper Confidence Bound
D2D	Device to Device
e.g.	For Example
FDD	Frequency Division Duplex

List of Abbreviations

HetNet	Heterogeneous Networks
ICI	Inter-Cell Interference
i.e.	That is
i.i.d.	Independent and Identically Distributed
KKT	Karush-Kuhn-Tucker
JT	Joint Transmission
LMI	Linear Matrix Inequality
LTE	Long Term Evolution
MAB	Multi-armed Bandit
MDP	Markov Decision Process
MISO	Multiple-Input Single-Output
MIMO	Multiple-Input Multiple-Output
MMSE	Minimum Mean Squared Error
mmWave	Millimeter Wave
NP	Non-deterministic Polynomial-time
OPEX	Operational Expenditure Cost
QoS	Quality of Service
RAN	Radio Access Network
RF	Radio Frequency
TDD	Time Division Duplex
UCB	Upper Confidence Bound
UT	User Terminal
SDP	Semidefinite Programming
SDR	Semidefinite Relaxation
SINR	Signal-to-Interference-plus-Noise Ratio
s.t.	Subject to
SWIPT	Simultaneous Wireless Information and Power Transfer
ZMCSCG	Zero-mean Circularly Symmetric Complex Gaussian

List of Notations

Throughout this thesis, the symbols are defined as follows:

w	A Scalar w
$\mathbf{w}$	A Vector $\mathbf{w}$
$\mathbf{W}$	A Matrix $\mathbf{W}$
$\mathbf{W} \succeq 0$	$\mathbf{W}$ is a Positive Semi-Definite Matrix
$\mathbf{W} \succ 0$	$\mathbf{W}$ is a Semi-Definite Matrix
Eigval($\mathbf{W}$)	Eigenvalue Operation on a Rank-one Matrix $\mathbf{W}$
Eigvec($\mathbf{W}$)	Eigenvector Operation on a Rank-one Matrix $\mathbf{W}$
$[\mathbf{W}]_{mn}$	the (m,n)-th Entry of $\mathbf{W}$
$\lvert w \rvert$	Magnitude of w
$\lVert \mathbf{w} \rVert$	Euclidean Norm of a Complex Vector $\mathbf{w}$
$\lVert \mathbf{w} \rVert^2$	Entry-Wise Absolute Value Square of a Complex Vector $\mathbf{w}$
$\mathbf{W}^T$	Transpose of $\mathbf{W}$
$\mathbf{W}^H$	Conjugate Transpose of $\mathbf{W}$
diag($\mathbf{w}$)	A Diagonal Matrix with Vector $\mathbf{w}$ on its Main Diagonal Entries
vec($\mathbf{W}$)	A Vector Obtained by Stacking $\mathbf{W}$ Columnwise
tr($\mathbf{W}$)	Trace of the Square Matrix $\mathbf{W}$
rank($\mathbf{W}$)	Rank of $\mathbf{W}$
$\mathbb{E}(.)$	Statistical Expectation for a Random Variable
$\mathbb{N}(.)$	Real Gaussian Random Variables
$\mathbb{CN}(\mu, \sigma^2)$	Complex Gaussian Random Variables with Mean μ and Variance σ^2
$\Re\{.\}$ and $\Im\{.\}$	Real and Imaginary Parts of the Argument, respectively

$\boldsymbol{A} \cdot \boldsymbol{B}$	Product of Two Matrices
$[\mathbf{A}\|\mathbf{B}]$	Concatenation of Matrices $\mathbf{A}$ and $\mathbf{B}$
$\Pr(.)$	Probability of an Input Random Event
$(\cdot)^*$	Optimal Solution of the Input Optimization Variable
$\max(x, y)$	Maximum Value between x and y
$[x]^+$	$\max(0, x)$
$[x, y]^-$	$\min(x, y)$
$\|.\|_0$	ℓ_0-norm indicating Number of Non-zero Entries in a Vector
$\|.\|_p$	ℓ_p-norm of a Vector
$O(.)$	An Upper Bound on a Function
$\mathbb{H}^{n\times m}$	Space of n-by-m Complex Hermitian Matrices
$\mathbb{C}^{n\times m}$	Space of n-by-m Complex Matrices
$\mathbb{R}^{n\times m}$	Space of n-by-m Real Matrices
$\mathbf{I}_n$	A n-by-n Identity Matrix
$\mathbf{0}_{m\times n}$	A m-by-n All-Zero Matrix
$\mathbf{1}_k$	A Column Vector with a One at the k-th Entry and Zeros Elsewhere with Appropriate Dimensionality

List of Nomenclature

$\mathcal{L}_b$	Index of N number of base stations (BSs)
$\mathcal{L}_i$	Index of K_i number of user terminals (UTs)
M	An array of M antenna elements per BS
UT_{ik}, $k \in \mathcal{L}_i$	the k-th single-antenna UT in the i-th cell
$\mathbf{w}_{ik} \in \mathbb{C}^{M\times 1}$	Beamforming vector from BS_i to UT_{ik}
$\mathbf{h}_{ijk} \in \mathbb{C}^{M\times 1}$	Channel vector from BS_i to UT_{jk}
s_{ik}	Data symbol for UT_{ik}
$n_{ik} \sim \mathbb{CN}(0, \sigma_{ik}^2)$	Zero-mean circularly symmetric complex Gaussian noise at UT_{ik}, with noise variance of σ_{ik}^2
$\rho_{ik} \in (0, 1)$	Maximum SINR outage probability
p_{ijk}	Intercell interference from BS i to UT_{jk}
γ_{ik}	Target SINR requested by UT_{ik}
$\hat{\mathbf{C}}_{ijk} = \mathbb{E}(\hat{\mathbf{h}}_{ijk}\hat{\mathbf{h}}_{ijk}^H)$	Estimated channel covariance matrix of UT_{jk}, as seen by BS_i
$\mathbf{\Delta}_{ijk} \in \mathbb{C}^{M\times M}$	Corresponding error matrix
$\mathbf{p}_i \in \mathbb{R}^{NK\times 1}$	Local intercell coupling variables at the i-th BS
$\mathbf{p} \in \mathbb{R}^{(N(N-1)+1)K\times 1}$	Global intercell coupling variables among all BSs
$\mathbf{X}_i \in \{0, 1\}$	A direction matrix to extract $\mathbf{p}_i$ from $\mathbf{p}$, i.e., $\mathbf{p}_i = \mathbf{X}_i\mathbf{p}$
$\mathbf{d}_{iik} \in \{0, 1\}$	A direction vector to extract $\sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik}$ from $\mathbf{p}$
$\delta = \lambda/2$	Spacing between two adjacent antenna elements
G_a	Array antenna gain
σ_a	Angular offset standard deviation
σ_s	Log-normal shadowing standard deviation

σ_F^2	Variance of the complex Gaussian fading coefficient
$L_{ijk}(\ell)$	Path loss model over a distance of ℓ m between BS_i and UT_{jk}
θ_{ijk}	Angle of departure for UT_{jk} with respect to the broadside of the antenna of BS_i
c_i	Percentage coefficient of the desired SINR targets that can be satisfied as a result of strict power constraints at BS i
P_i	Transmit power restriction of the i-th BS
η_{ik}	SINR satisfaction ratio of the achieved SINR over the scaled target SINR of UT_{ik}
$\mathcal{T}$	Index of T number of discrete time slots
$\mathcal{K}$	Index of K number of learning trials within a time slot
$\mathcal{P}$	Index of P number of periods
$E_n^{[\text{a}]}(t)$	Amount of ahead-of-time purchased energy (base arm) at BS n at the t-th time slot
$\pi^{[\text{a}]}$	Per unit energy price of $E_n^{[\text{a}]}(t)$
$E_n^{[\text{r}]}(t)$	Amount of real-time energy shortage to be supplied by the spot market in the grid for the n-th BS at the t-th time slot
$\pi^{[\text{r}]}$	Per unit energy price of $E_n^{[\text{r}]}(t)$
$S_n(t)$	Amount of surplus energy of the n-th BS at the t-th time slot to be sold back to the grid
$\pi^{[\text{e}]}$	Per unit energy price of $S_n(t)$
$G_n(t)$	Amount of renewable energy generation of BS n at time slot t
$\pi^{[\text{g}]}$	Per unit equivalent annual cost of renewable harvesters for $G_n(t)$
$C^{[\text{total}]}(t)$	Total energy cost incurred by the n-th BS at the t-th time slot
$P_n^{[\text{Tx}]}(t)$	Total transmit power at the n-th BS at the t-th time slot
$P_n^{[\text{c}]}$	Hardware circuit power consumption at the n-th BS
$P_n^{[\text{total}]}(t)$	Total energy consumption of the n-th BS at the t-th time slot
η	Power amplifier efficiency
$P_n^{[\text{Tmax}]}$	Maximum transmit power allowance at the n-th BS

$B_n^{[\text{fronthaul}]}(t)$	Fronthaul capacity consumption of BS n at time slot t
$B_n^{[\text{limit}]}$	Fronthaul capacity limit of BS n
$R_i(t)$	Achievable data rate (bit/s/Hz) for UT i at time slot t
ξ_{ni}	Weight factor for sparse beamforming
$\mathcal{J}$	Index of J number of discrete ahead-of-time energy packages (base arms) $\{\mathcal{E}^1, \cdots, \mathcal{E}^J\}$ offered by the grid
$\mathcal{S}^{[\text{set}]}(k)$	N ahead-of-time energy packages (super arm) at the k-th trial for the next time slot
$\mathcal{R}(E_n^{[\text{a}]}(k))$	Individual reward for base arm $E_n^{[\text{a}]}(k)$ at the k-th trial
$\mathbf{r_n^{[k,t]}}$	Reward vector of the n-th BS
$r_{n,e}^{[k,t]}, e \in \mathcal{J}$	Reward associated to the e-th base arm in trial k at time slot t
$\hat{\mathbf{r}}_\mathbf{n}^{[\mathbf{t}]}$	Estimated mean reward vector for BS n at time slot t
$\bar{\mathbf{r}}_\mathbf{n}^{[\mathbf{t}]}$	Adjusted reward vector for BS n at time slot t
$E_n^{[\text{s}]}(t)$	Amount of initial energy contents of the storage of BS n in the beginning of the t-th time slot
$\pi^{[\text{s}]}$	Per unit equivalent annual cost of storage devices for $E_n^{[\text{s}]}(t)$
$E_n^{[\text{c}]}(t)$	Amount of energy (base arm) charged to the storage of BS n prior to the actual time of energy demand at the t-th time slot
$\pi^{[\text{c}]}$	Per unit energy price of $E_n^{[\text{c}]}(t)$
$E_n^{[\text{capacity}]}(t)$	Upper limit of the storage capacity of BS n
$\mathcal{R}_t(E_n^{[\text{c}]}(k))$	Reward of the arm selected for the n-th BS at the t-th time slot
$\mathcal{D}$	Discount factor indicating the importance of previous rewards

Chapter 1

Introduction

1.1 Motivation

Last decade has witnessed the evolution of information and communication technology incurred by the explosive increase in the number of mobile subscribers and the serving wireless devices. By 2030, the number of wireless devices will rise to 100 billion and result in extremely massive connectivity [13]. Such massive connectivity requires enormous energy consumption in mobile communications networks and with current technique, can only be achieved at the expense of incredible greenhouse gas emissions. Current network operators relying on the fossil-fuel-based electric energy generation to power their networks, contribute to a significant proportion to the global carbon footprint, with a share of approximately 2 percent [13]. Vodafone, for instance, used more than 1 million gallons of diesel per day in 2011 to power their network [14]. The amount of carbon dioxide (CO2) equivalent emissions of cellular networks are, respectively, 86 and 170 million tons for 3G and 4G networks [15]. Such amount of emissions are expected to escalate to 345 million tons by 2020 [16], indicating a steeper rise as compared to the prediction of 235 million tons in the SMART2020 report in 2008 [17]. In order to support the ever-increasing demand for high data rate communications with seamless coverage and diverse quality of service (QoS), the 5G networks are expected to be launched by 2020 and provide more than 1000 times the system capacity as well as 10 times the energy efficiency of the 4G networks [18], as shown in Fig. 1.1, which raises numerous challenges

to be addressed by the research community. Amongst which, the enormous energy

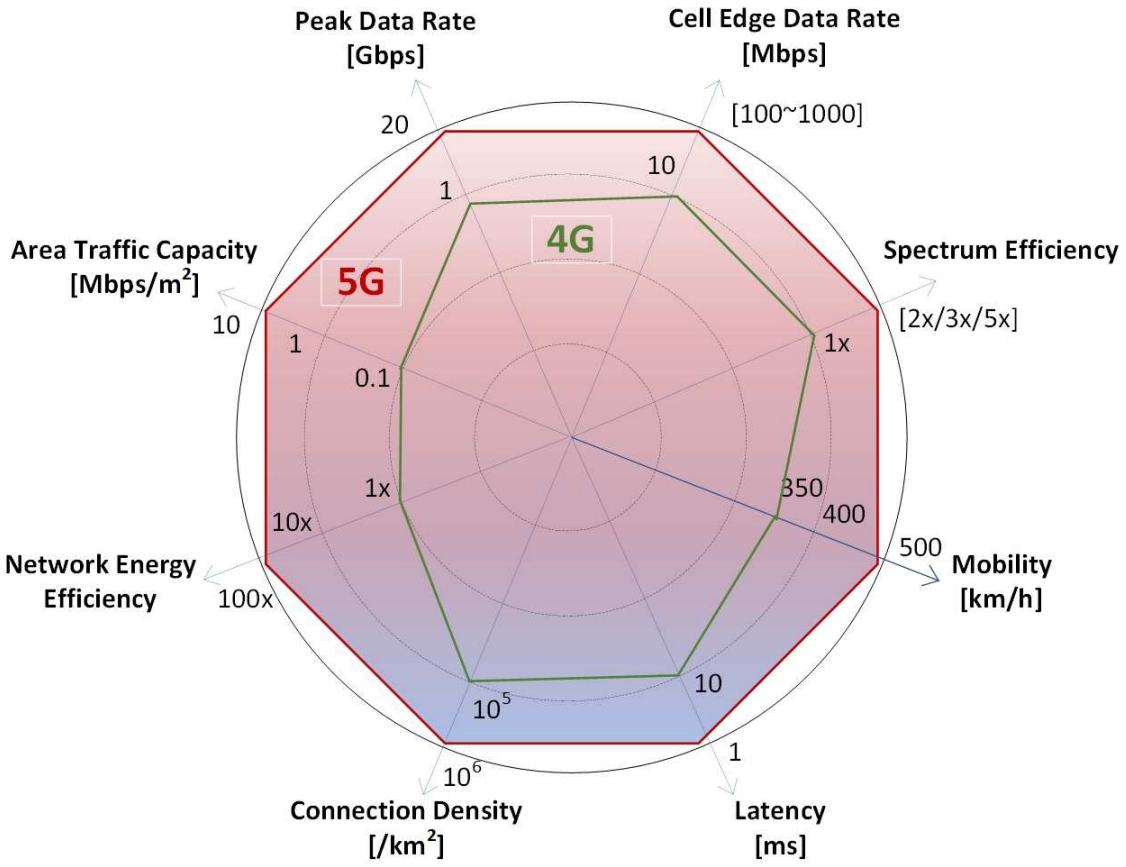


Figure 1.1: 5G technical improvement over 4G.

consumption arose by next generation dense wireless communication networks with millions more BSs and billions of connected devices, has always been considered as one of the most challenging issues from both ecological and economic perspectives. It has been revealed that approximately 30 percent of total operational expenditure (OPEX) of mobile network operators is energy cost [13], whilst the energy consumption of base stations (BSs) contributes to more than 70 percent of operators' total electricity bill [18]. The energy consumption of a BS consists of the radiated energy, the energy loss due to efficiency of the non-ideal power amplifiers and static energy dissipated in all other hardware blocks of the transmit-receive chains, e.g., A/D conversion, filtering and cooling operations. It is usually assumed in the literature that the transmit

amplifiers operate in the linear region, and the static hardware energy is independent of the radiated energy [19]. This coupled with the long-standing resource scarcity in mobile networks caused by the mounting growth of mobile subscribers and energy demand motivate the research interests in energy-efficient wireless communication (also known as green communication) in the recent years, where the radiated energy consumption has become a primary concern in the design and operation of wireless communication networks.

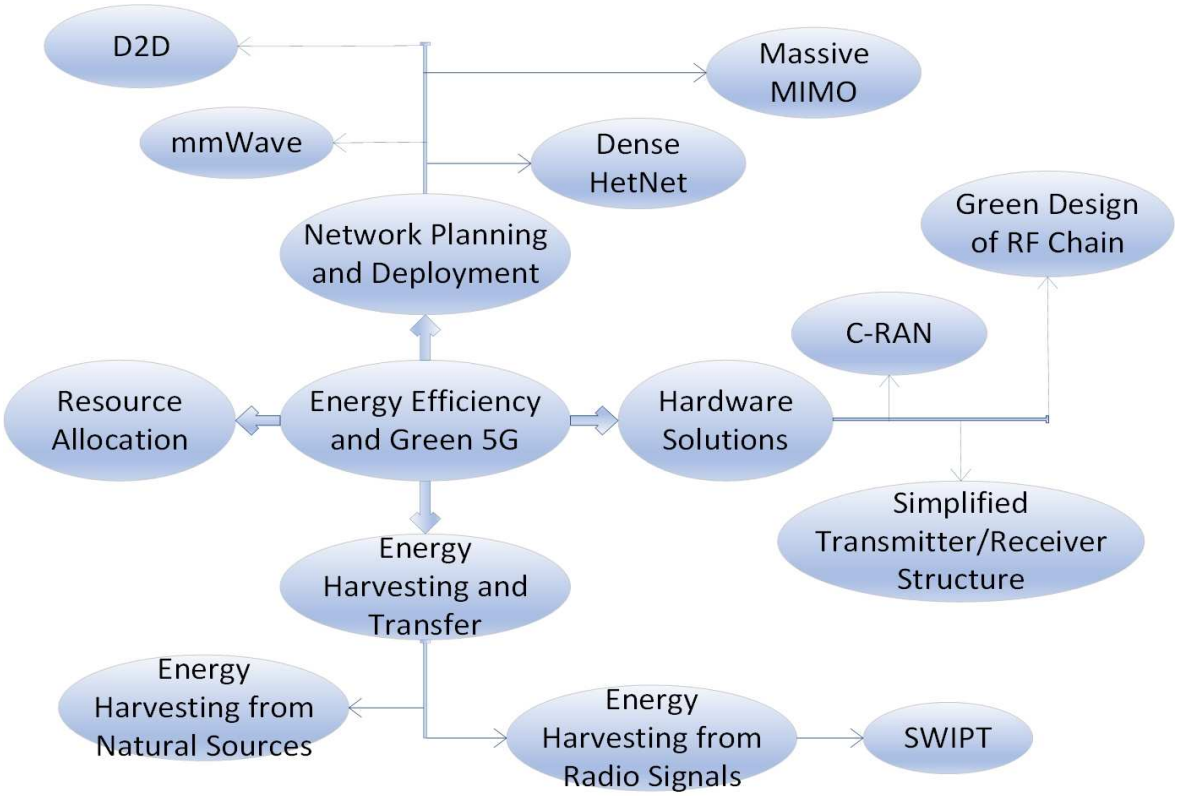


Figure 1.2: Energy efficient 5G solutions.

Numerous proposals and research projects have been launched around the world to reduce substantially the total energy consumption for the entire radio access network through various energy-efficient techniques [14]. As illustrated in Fig. 1.2, energy efficient techniques can be classified into four main categories: resource allocation, network planning and deployment, energy harvesting and hardware solutions [19], where the state-of-the-art research focuses are briefly presented as

follows.

- Resource allocation technique increases energy efficiency by allocating the system radio resources to maximize the amount of information that is reliably transmitted per Joule of consumed energy.
- Massive multiple-input multiple-output (MIMO) has shown its ability in reducing the radiated energy and averaging out multi-user interference, provided that the favorable propagation condition holds. Whereas, the practical deployment and hardware impairment are considered as major challenges.
- mmWave increases network bandwidth and offloads traffic from the sub-6GHz cellular frequencies for short range dense communications. However, the implementation of digital beamforming raises complexity, energy consumption and cost issues.
- Simultaneous wireless information and power transfer (SWIPT) has been widely investigated in the literature for sustainable operation of battery limited devices by exploiting signals transmitted from BSs [4, 19].
- Dense heterogeneous networks (HetNet) reduces the distances between nodes and UTs, thus provides higher data rates at lower power consumption, provided that a balance between density level and interference control is achieved.
- Cloud radio access network (C-RAN) that detaches baseband processing units from conventional BSs and groups them in a pool, not only enables the potentiality and flexibility of mobile edge computing in the network, but also provides substantial energy and deployment cost savings.

Due to resource scarcity of cellular networks, small cells has become one of the research focuses recently and has been considered as a promising method to expand service coverage and increase network capacity at an attractive cost for future ultra-dense heterogeneous networks, as shown in Fig. 1.3. Small cells are low-powered radio access nodes, e.g., microcells, picocells and femtocells, that are "small" in terms

of capability and coverage area as compared to macrocells and are easy to deploy and maintain. They make best use of available spectrum by reusing the same frequencies many times within a geographical area. Beamforming technique for directional signal transmission and reception, can further enhance small cell coverage. Recently, the

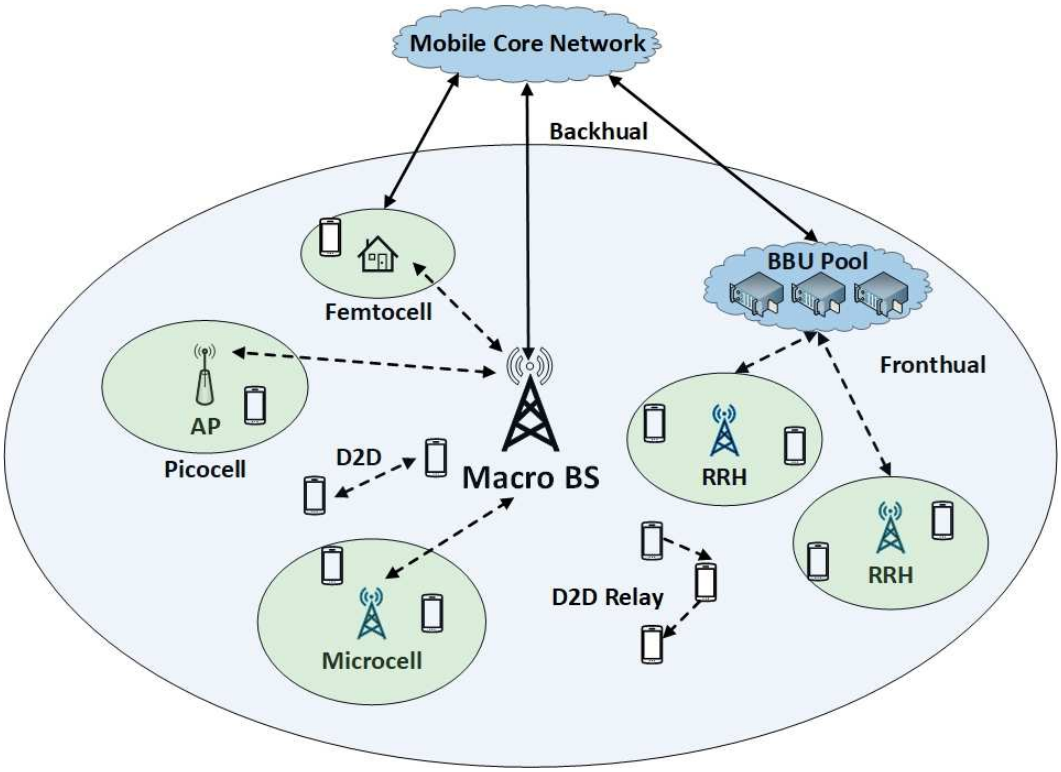


Figure 1.3: An example of 5G dense heterogeneous network.

orthogonal frequency-division multiple access based systems such as Long Term Evolution (LTE) are being deployed with a frequency reuse of 1, i.e., operated under a shared bandwidth. However, the resulting intercell interference (ICI), i.e., the signals at the same frequency received by user terminals (UTs) from undesired transmitters, may lead to significant system performance degradation. Consequently, interference coordination is a key factor for minimizing energy consumption in green communications and hence, reducing total energy cost of the network operators. In recent years, coordinated transmission, where multiple BSs collaborate at either signal level or beamforming level to serve individual UTs, has been recognized as a key enabling technique to mitigate ICI and substantially improve system capacity

[20]. This scenario, nevertheless, requires all information to be circulated among BSs, which may be infeasible for practical capacity-constrained fronthaul links. Consequently, sparse beamforming technique for partial BS cooperation as well as coordinated transmission in a distributed manner, have attracted the attention of researchers in recent years [6, 11, 21].

Apart from energy efficiency and interference control for green communications [22], powering the BSs with renewable energy generated from naturally replenished environmental resources ranging from sunlight, wind, tides and waves to geothermal heat [23], has also been regarded as a promising technology for the next generation green wireless networks from ecological economics perspective. The network operators relying on the conventional fossil-fuel-based electric grid not only face potential challenge of drastically increased operational costs due to growing amount of on-grid energy consumption for future dense networks, but also significantly contribute to the global footprint [15]. Powering BSs with renewable energy is beneficial in terms of not only reducing greenhouse gas emissions, but also the potential of reducing the energy cost of the network/grid operators, since the cost of renewable energy generation in general is much lower than that of the energy from the conventional grid [24]. The renewable energy in 2010 contributed only 16.6 per cent to the total energy generation of Europe Union, whilst by 2040, its contribution is expected to reach 47.7 percent [25]. However, the renewable energy generation is naturally uncontrollable and non-dispatchable since it highly depends on location, time and efficiency of the harvesting devices, e.g., solar panels and wind turbines [26]. Thus, in some circumstances, it is insufficient to meet the energy demand of the networks. Realizing these features and providing the opportunity to the reliable and cost-efficient operation of wireless networks motivate the integration of the renewable energy with the conventional electric grid to power next generation wireless networks. The solutions to the wireless channel random dynamism as well as the intermittent nature of renewable energy supply and significant electricity price fluctuation, meanwhile, are currently of great interest for the research community

[7, 12, 27, 28].

This thesis investigates energy management and interference control for green communications in multi-cell interference networks from both ecological and economic perspectives. From the first perspective, the cross-link coupling effect among a cluster of BSs, e.g., ICI, is taken into account and the alternatives to the existing coordinated transmission strategies for further reduction of energy consumption as well as robustness against the imperfect channel state information (CSI) are examined. From the second perspective, novel reinforcement learning based algorithms that adapt to the dynamic nature of wireless networks as well as renewable energy supply are developed to achieve a reliable and cost-efficient operation of the BSs supplied by a hybrid grid/renewable energy generators. The scope of research on energy management strategies carried out in this thesis is shown in Fig. 1.4.

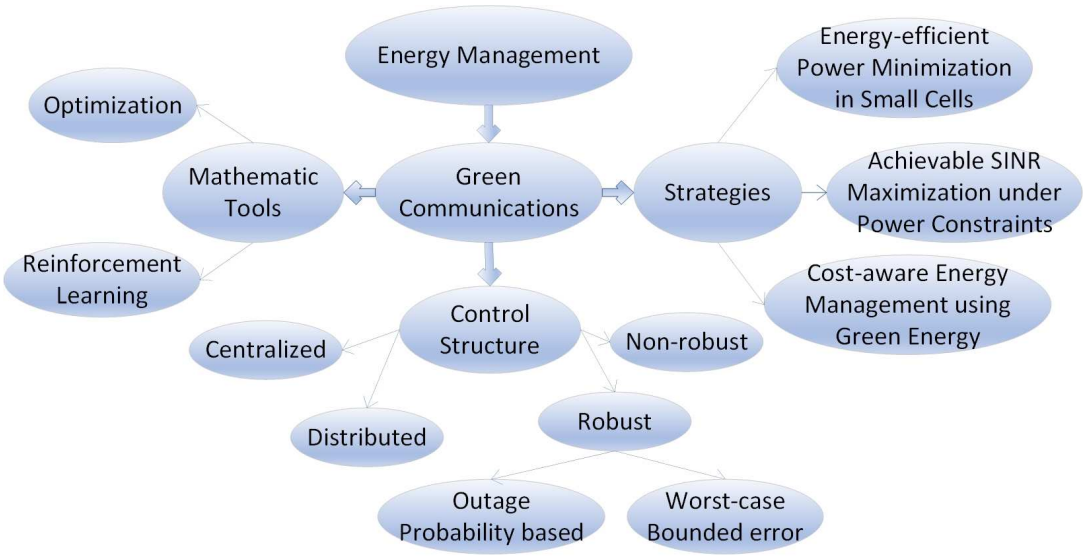


Figure 1.4: Scope of research on energy management in this thesis.

1.2 Contributions of the Thesis

Addressing the problems stated in Section 1.1, this thesis contributes to the learning-based energy management and interference control among BSs for green

communications in multi-cell networks. The major contributions of this thesis are summarized as follows.

- Chapter 3 considers the robust optimization problem of minimizing the aggregate downlink transmit power in a distributed manner in the presence of imperfect CSI in multicell interference networks. Due to the fact that worst-case is a rare occurrence in practical network, this problem is constrained to satisfying a set of signal-to-interference-plus-noise-ratio (SINR) requirements and providing robustness against the second order statistical and instantaneous CSI uncertainties at individual UTs with certain SINR outage probabilities. Taking into account the cross-link coupling effect among a cluster of BSs, the individual BSs gradually learn the ICI imposed by other BSs via subgradient iterations with a light inter-BS communication overhead. These contributions have been published in [29] and [30].

- Chapter 4 investigates the problem of maximizing the weighted SINR requirements at UTs in a distributed manner in multicell interference networks. The optimization is constrained to strict individual BS transmit power limitations in the presence of imperfect CSI in the worst-case scenario. Instead of solving the optimization problem directly, the original problem is converted into an equivalent total transmit power minimization problem. An upper confidence bound based algorithm is proposed for the individual BSs to distributively learn the optimal achievable percentage coefficient of SINR targets based on per BS power budget and coordinate ICI across the BSs via light inter-BS communications. This contribution has been published in [31].

- Chapter 5 introduces a reinforcement learning algorithm inspired by combinatorial multi-armed bandit (CMAB) model to minimize the time-averaged energy cost at individual BSs, powered by various energy markets and local renewable energy sources, over a finite time horizon. The algorithm sustains traffic demands by enabling sparse beamforming to

schedule dynamic user-to-BS allocation and proactive energy provisioning at BSs to make ahead-of-time price-aware energy management decisions. This contribution has been published in [32].

- To address the dynamic statistics of wireless networks as well as the variability of renewable energy supply and energy prices that are practically unknown in advance, an adaptive CMAB-inspired strategy for energy storage management and cost-aware coordinated load control at the BSs is developed in Chapter 6. The proposed strategy makes online foresighted decisions on the amount of energy to be stored in storage to minimize the average energy cost over long time horizon. This contribution has been published in [33].

1.3 Outline of the Thesis

The thesis is organized as follows. Chapter 1 introduces the motivation and contributions of the thesis. In Chapter 2, mathematical preliminaries such as convex optimization are introduced, followed by some basic concepts and literature survey of the state-of-the-art techniques in cooperative transmission, energy management and reinforcement learning. The major contributions of this thesis are included in Chapter 3,4,5 and 6. More specifically, Chapter 3 investigates the robust distributed ICI coordination and power minimization problem constrained by certain SINR outage probabilities via sub-gradient iterative ICI learning among BSs. Then, the problem of maximizing the weighted SINR requirements is studied in Chapter 4 and an upper confidence bound inspired learning algorithm is proposed to optimally and distributively coordinate ICI and scale the SINR targets based on per-BS power budget. Chapter 5 and 6, on the other hand, address the dynamic environment statistics and focus on CMAB-inspired adaptive online learning algorithms for cost-aware energy management of BSs deployed with and without storage units, respectively. Finally, Chapter 7 summarizes the thesis and indicates some interesting future research directions.

Chapter 2

Background

2.1 Introduction

This chapter aims to provide a general overview of downlink energy management in multi-cell interference networks as well as mathematical preliminaries that will be used in the subsequent chapters. In this chapter, the mathematical preliminaries such as convex optimization will be firstly introduced, followed by some basic concepts and literature review of the state-of-the-art robust downlink cooperative transmission strategies, energy management designs and recent advances in reinforcement learning in wireless communication networks.

2.2 Mathematical Preliminaries

2.2.1 Convex Optimization

Convex optimization is a subfield of optimization that seeks the minimum of convex functions over convex sets. One key advantage of convex problems over non-convex optimization problems is that the convex problems can be solved efficiently using powerful numerical algorithms even when a closed form does no exist. Due to the convexity of both objective functions and convex constraints, a local minimum in a convex optimization problem must be a global minimum, and there exists a rigorous optimality condition as well as a duality theory to

verify the optimal solution [34]. In addition, convexity provides possibility to address the difficult non-convex problems using convex approximations that are more efficient than classical linear ones. The convex optimization has been widely applied to provide efficient and reliable solutions to large practical engineering application problems in various disciplines such as automatic control systems, signal processing, communications and networks. Many communication problems can either be formulated as or converted into convex optimization problems to greatly facilitate their analytic and numerical solutions [34].

As illustrated in Fig. 2.1, a set C is convex if for $\forall x, y \in C$, we have

$$\theta x + (1-\theta)y \in C, \quad \forall \theta \in [0,1], \tag{2.1}$$

which indicates that the line segment between any two point $x, y \in C$ lies in C.

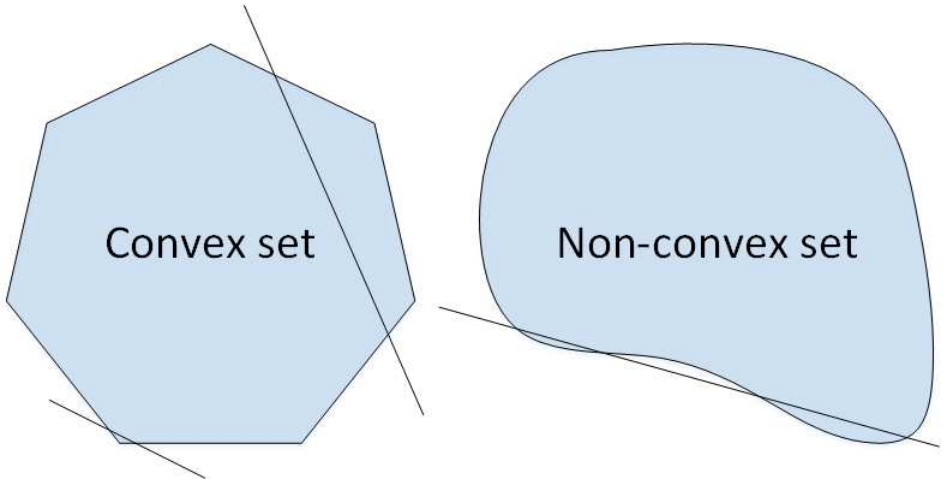


Figure 2.1: Some simple convex and non-convex sets.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if **dom** f is a convex set and for $\forall x, y \in$ **dom** f, we have

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y), \quad \forall \theta \in [0,1], \tag{2.2}$$

which indicates that the line segment between $(x, f(x))$ and $(y, f(y))$ always

dominates the function f, as indicated by Fig. 2.2 [8]. If strict inequality holds in (2.2) for $\forall x \neq y,\ \theta \in (0,1)$, the function f is said to be strictly convex. Moreover, f is said to be concave if $-f$ is convex, and strictly concave if $-f$ is strictly convex.

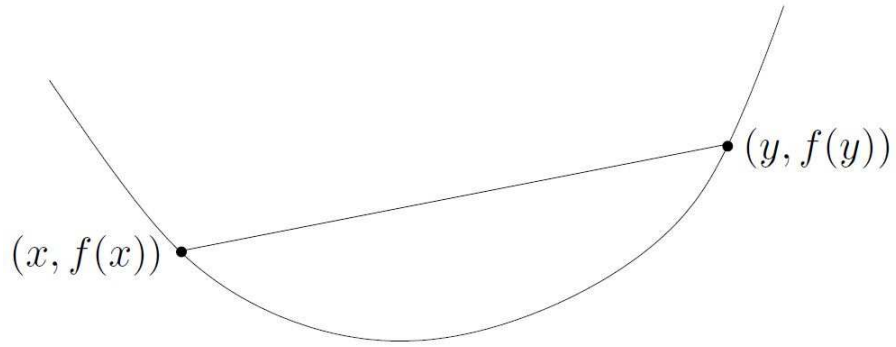


Figure 2.2: Example of a convex function [8].

The general form of convex optimization problem [8] can be expressed as follows,

$$\begin{aligned} \min_{x} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m, \\ & h_i(x) = 0, \quad i = 1, \cdots, p, \end{aligned} \tag{2.3}$$

where $x \in \mathbb{R}^n$, $f_0 : \mathbb{R}^n \to \mathbb{R}$, $f_i : \mathbb{R}^n \to \mathbb{R}$ and $h_i : \mathbb{R}^n \to \mathbb{R}$, respectively, are defined as the optimization variable, the objective function, the inequality constraint functions and the equality constraint functions. The objective is to find an x that minimizes $f_0(x)$ among all x that satisfy the conditions $f_i(x) \leq 0, i = 1, \cdots, m$ and $h_i(x) = 0, i = 1, \cdots, p$. By definition, the functions $h_i : \mathbb{R}^n \to \mathbb{R}$ are affine and the functions $f_0, \cdots, f_i : \mathbb{R}^n \to \mathbb{R}$ are convex, i.e., satisfy

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y) \tag{2.4}$$

for all $x, y \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ with $\alpha + \beta = 1, \alpha \geq 0, \beta \geq 0$.

The problem (2.3) is said to be feasible if there exists at least one feasible point $x \in \mathcal{D}$ that satisfies constraints $f_i(x) \leq 0,\ i = 1, \cdots, m$ and $h_i(x) = 0,\ i = 1, \cdots, p$,

where the domain of problem (2.3), i.e.,

$$\mathcal{D} = \bigcap_{i=0}^{m} \text{dom } f_i \cap \bigcap_{i=0}^{p} \text{dom } h_i \tag{2.5}$$

is a convex set. The optimal value p^* of problem in (2.3) is defined as

$$p^* = \inf \{f_0(x) \mid f_i(x) \leq 0,\ i = 1, \cdots, m, h_i(x) = 0,\ i = 1, \cdots, p\}, \tag{2.6}$$

where $p^* = \infty$ if the problem is infeasible and $p^* = -\infty$ if the problem is unbounded below. In addition, x^* is said to be an optimal point if x^* is feasible and $f_0(x^*) = p^*$, whilst a feasible point x is said to be ϵ-suboptimal if $f_0(x) \leq p^* + \epsilon$.

The introduction of slack variables that replace each inequality constraint with an equality constraint and a nonnegativity constraint, is commonly used in transformations of convex optimization problems. Introducing a new variable $s_i \in \mathbb{R}^m$, the convex problem in (2.3) can be transformed into an equivalent problem, as

$$\begin{aligned} \min_{x} \quad & f_0(x) \\ \text{subject to} \quad & s_i \geq 0, \quad i = 1, \cdots, m, \\ & f_i(x) + s_i = 0, \quad i = 1, \cdots, m, \\ & h_i(x) = 0, \quad i = 1, \cdots, p. \end{aligned} \tag{2.7}$$

In general, there is no analytical formula but effective methods for the solution of the convex optimization problems. With recent development in optimization theory, e.g., interior-point method, solving convex optimization problem such as semidefinite programming (SDP) is almost as straightforwardly as solving linear programming. For instance, the efficient and reliable interior-point methods [35] can be proved in some practical cases, to solve the convex optimization problem to a specified accuracy with a number of operations that does not exceed a polynomial of the problem dimensions. The convex optimization problem in (2.3) ordinarily can be

solved by the interior-point methods in a number of steps ranging between 10 and 100, and each step requires on the order of $\max\{n^3, n^2m, F\}$ operations, where F is the cost of evaluating the first and second derivatives of $f0, \cdots, fm$ [8]. In practice, modern solvers for solving convex optimization problems, e.g., the SeDuMi solver [36], either generate an optimal solution or an indication of infeasibility. Solvers for non-convex optimization problems, nevertheless, typically fail to converge when the underlying problem is infeasible, due to either data overflow or the exceeding of maximum number of iterations [34].

2.2.2 Lagrangian Duality and Karush-Kuhn-Tucker Condition

Consider problem (2.3) as the primal optimization problem and x as the primal vector, the Lagrangian duality is to augment the objective function with a weighted sum of the constraints functions in (2.3) [8]. The Lagrangian function $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ associated with the problem in (2.3) can be formulated as

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \tag{2.8}$$

with dom $L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$, where $\lambda \in \mathbb{R}^m$ and $\nu \in \mathbb{R}^p$, respectively, are the lagrange multiplier vectors associated with the inequality constraints $f_i(x) \leq 0,\ i = 1, \cdots, m$ and the equality constraints $h_i(x) = 0,\ i = 1, \cdots, p$ [34].

The dual function $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ associated with problem (2.3) is defined as

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} (f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)), \tag{2.9}$$

and as a pointwise infimum of a family of affine functions of (λ, ν), it is always concave [34].

Let us consider the convex primal problem in (2.3), the following Karush-Kuhn-Tucker (KKT) conditions are not only necessary but also sufficient

for the points $\tilde{x}$ and $(\tilde{\lambda}, \tilde{\mu})$ to be primal and dual optimal, with zero duality gap.

$$\begin{aligned} f_i(\tilde{x}) &\leq 0, \quad i = 1, \cdots, m, \\ h_i(\tilde{x}) &= 0, \quad i = 1, \cdots, p, \\ \tilde{\lambda}_i &\geq 0, \quad i = 1, \cdots, m, \\ \tilde{\lambda}_i f_i(\tilde{x}) &= 0, \quad i = 1, \cdots, m, \\ \nabla f_0(\tilde{x}) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla f_i(\tilde{x}) + \sum_{i=1}^{p} \tilde{\mu}_i \nabla h_i(\tilde{x}) &= 0. \end{aligned} \tag{2.10}$$

2.2.3 Semidefinite Programming

SDP is a relatively new subfield of convex optimization that minimizes a linear objective function over the intersection of the cone of positive semidefinite matrices with an affine space, where the affine constraints include both equalities and inequalities. Many practical problems can be modeled or approximated as SDP problems and efficiently solved by interior point methods. The SeDuMi solver [36] introduced in the previous section, is commonly used to solve SDP problems, whilst the CVX [35] that supports SeDuMi solver, will be adopted to solve the SDP problems with linear matrix inequality (LMI) constraints in the following chapters. A standard SDP has the following form with LMI constraints, as

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{subject to} \quad & F(x) \succeq 0, \end{aligned} \tag{2.11}$$

where $x \in \mathbb{R}^n$ is the optimization variable and LMI constraints, i.e.,

$$F(x) \triangleq F_0 + \sum_{i=1}^{n} x_i F_i \succeq 0, \quad F_0, \cdots, F_n \in \mathbb{R}^{m \times m} \tag{2.12}$$

is a Hermitian matrix and indicates that $F(x)$ is positive semidefinite, i.e., $z^T F(x) z \geq 0, \ \forall z \in \mathbb{R}^m$ [37].

2.2.4 Multi-cell Multi-user Downlink Beamforming

Consider N base stations (BSs) equipped with M antennas jointly design their beamforming vectors and transmit to their K local single-antenna user terminals (UTs) over a shared bandwidth. Then, the signal received by the k-th UT in the i-th cell can be expressed as

$$z_{ik} = \mathbf{h}_{iik}^H \mathbf{w}_{ik} s_{ik} + \sum_{n \neq k}^{K} \mathbf{h}_{iik}^H \mathbf{w}_{in} s_{in} + \sum_{j \neq i}^{N} \sum_{m=1}^{K} \mathbf{h}_{jik}^H \mathbf{w}_{jm} s_{jm} + n_{ik}, \tag{2.13}$$

where s_{ik} is the data symbol for UT_{ik}, $\mathbf{w}_{ik} \in \mathbb{C}^{M \times 1}$ denotes the beamforming vector from BS i to UT k, $n_{ik} \sim \mathbb{CN}(0, \sigma_{ik}^2)$ is the additive white Gaussian noise and $\mathbf{h}_{ijk} \in \mathbb{C}^{M \times 1}$ indicates the channel vector from BS_i to UT k in cell j. Without loss of generality, let $\mathbb{E}(|s_{ik}|^2) = 1$, then the signal-to-interference-plus-noise-ratio (SINR) at the k-th UT in cell i is given by

$$\mathrm{SINR}_{ik} = \frac{|\mathbf{h}_{iik}^H \mathbf{w}_{ik}|^2}{\sum_{n \neq k}^{K} |\mathbf{h}_{iik}^H \mathbf{w}_{in}|^2 + \sum_{j \neq i}^{N} \sum_{m=1}^{K} |\mathbf{h}_{jik}^H \mathbf{w}_{jm}|^2 + \sigma_{ik}^2}. \tag{2.14}$$

A typical downlink beamforming design is to find the optimal set of $\mathbf{w}_{ik}$ that minimizes the overall transmit power while guaranteeing the SINR requirements γ_{ik} at the individual UTs, as

$$\begin{aligned} \min_{\mathbf{w}_{ik}, \forall i,k} \quad & \sum_{i=1}^{N} \sum_{k=1}^{K} \|\mathbf{w}_{ik}\|^2 \\ \text{s.t.} \quad & \frac{|\mathbf{h}_{iik}^H \mathbf{w}_{ik}|^2}{\sum_{n \neq k}^{K} |\mathbf{h}_{iik}^H \mathbf{w}_{in}|^2 + \sum_{j \neq i}^{N} \sum_{m=1}^{K} |\mathbf{h}_{jik}^H \mathbf{w}_{jm}|^2 + \sigma_{ik}^2} \geq \gamma_{ik}, \quad \forall i, k. \end{aligned} \tag{2.15}$$

It can be verified that the SINR constraints in (2.15) are non-convex. In the sequel, the problem in (2.15) will be transformed to a SDP form with LMI using the semidefinite relaxation (SDR) technique. The SDR is a powerful and efficient

approximation technique in the area of signal processing and communications, and has proved its capability of providing accurate and sometimes near-optimal approximation [38].

The problem in (2.15) can be converted to the following format

$$\begin{aligned} \min_{\mathbf{w}_{ik},\forall i,k} \quad & \sum_{i=1}^{N}\sum_{k=1}^{K}\mathbf{w}_{ik}^H\mathbf{w}_{ik} \\ \text{s.t.} \quad & \frac{\mathbf{h}_{iik}^H\mathbf{w}_{ik}\mathbf{w}_{ik}^H\mathbf{h}_{iik}}{\sum_{n\neq k}^{K}\mathbf{h}_{iik}^H\mathbf{w}_{in}\mathbf{w}_{in}^H\mathbf{h}_{iik}+\sum_{j\neq i}^{N}\sum_{m=1}^{K}\mathbf{h}_{jik}^H\mathbf{w}_{jm}\mathbf{w}_{jm}^H\mathbf{h}_{jik}+\sigma_{ik}^2} \geq \gamma_{ik}, \quad \forall i,k. \end{aligned} \tag{2.16}$$

Let us define $\mathbf{H}_{iik} = \mathbf{h}_{iik}\mathbf{h}_{iik}^H$ and $\mathbf{W}_{ik} = \mathbf{w}_{ik}\mathbf{w}_{ik}^H$. It is evident that $\mathbf{W}_{ik}$ is a semidefinite and Hermitian matrix with $\text{rank}(\mathbf{W}_{ik}) = 1$. Considering the following equality

$$\mathbf{w}_{ik}^H\mathbf{w}_{ik} = \text{tr}(\mathbf{w}_{ik}^H\mathbf{w}_{ik}) = \text{tr}(\mathbf{w}_{ik}\mathbf{w}_{ik}^H) = \text{tr}(\mathbf{W}_{ik}), \tag{2.17}$$

the problem in (2.15) can be transformed as

$$\begin{aligned} & \min_{\mathbf{W}_{ik},\forall i,k} \sum_{i=1}^{N}\sum_{k=1}^{K}\text{tr}(\mathbf{W}_{ik}) \\ \text{s.t.} \quad & \gamma_{ik}^{-1}\text{tr}(\mathbf{H}_{iik}\mathbf{W}_{ik}) \geq \sum_{n\neq k}^{K}\text{tr}(\mathbf{H}_{iik}\mathbf{W}_{in}) + \sum_{j\neq i}^{N}\sum_{m=1}^{K}\text{tr}(\mathbf{H}_{jik}\mathbf{W}_{jm}) + \sigma_{ik}^2, \\ & \mathbf{W}_{ik} = \mathbf{W}_{ik}^H \succeq 0, \\ & \text{rank}(\mathbf{W}_{ik}) = 1, \ \forall i,k. \end{aligned} \tag{2.18}$$

Note that relaxing the non-convex rank-one constraints of $\text{rank}(\mathbf{W}_{ik}) = 1$ in (2.18)

via SDR technique results in a SDP form, i.e.,

$$\min_{\mathbf{W}_{ik},\forall i,k} \sum_{i=1}^{N}\sum_{k=1}^{K} \mathrm{tr}(\mathbf{W}_{ik}) \tag{2.19}$$

$$\text{s.t.} \quad \gamma_{ik}^{-1}\mathrm{tr}(\mathbf{H}_{iik}\mathbf{W}_{ik}) - \sum_{n\neq k}^{K} \mathrm{tr}(\mathbf{H}_{iik}\mathbf{W}_{in}) - \sum_{j\neq i}^{N}\sum_{m=1}^{K} \mathrm{tr}(\mathbf{H}_{jik}\mathbf{W}_{jm}) - \sigma_{ik}^2 \geq 0,$$

$$\mathbf{W}_{ik} = \mathbf{W}_{ik}^H \succeq 0, \ \forall i,k.$$

For general non-convex quadratic problems, solving a relaxed SDP problem does not always yields the optimal rank-one solutions, it usually leads to conveniently tractable numerical solutions, i.e., optimal solutions with rank greater than one. In such cases, the SDR technique can only provide a lower bound on the optimal objective function and possibly attain an approximate solution to the original problem [39]. If the optimal solution $\mathbf{W}_{ik}^*$ to problem (2.19) is rank-one, the optimal beamformer $\mathbf{w}_{ik}^*$ is the eigenvector of $\mathbf{W}_{ik}^*$. Otherwise, the randomization and scaling procedures can be implemented by the BSs to get the near-optimal beamformers [3, 11, 40]. Note that with a sufficient number of trials of the randomization procedure, the gap between the optimal and the suboptimal values can be arbitrarily reduced. For more discussion on rank-one relaxation, please refer to [41] and the references therein.

2.3 Multi-cell Interference Network and Cooperative Transmission

Heterogeneous networks that consists of multiple types of transmission points to support various size of wireless coverage zones, has been regarded as a key enabling technology to support the ever-increasing mobile data traffic and high data rate communications with seamless and ubiquitous quality of service (QoS) for next generation wireless communication networks [42, 43]. Long Term Evolution (LTE)-Advanced multi-cell multiuser networks suffer from intra-cell interference as a result of simultaneously transmission to multiple users, as well as intercell interference

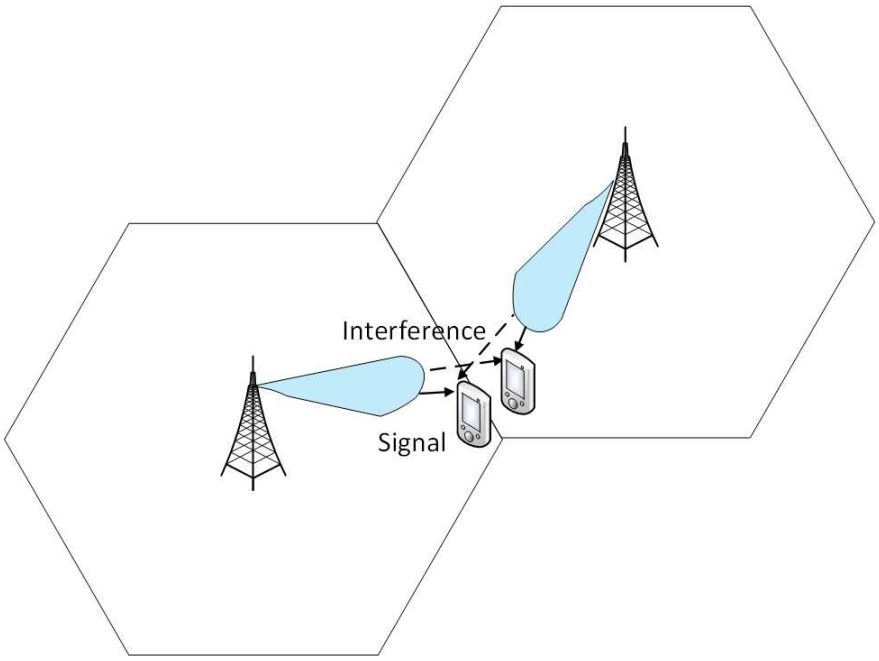


Figure 2.3: A typical example of multi-cell multiuser interference network.

(ICI) among neighboring cells as a consequence of the ever shrinking cell sizes [44], as illustrated in Fig. 2.3. Without proper interference control mechanisms, the system performance, especially at cell edge, can be degraded significantly. One possible solution is to equip BSs and/or UTs with multiple smart antennas and smart signal processing algorithms [45] to create different radiation patterns, and adopt beamforming techniques to separate the intended signal (beam) to the individual UTs as much as possible. Coordination techniques such as semi-static fractional frequency reuse and dynamic coordinated transmission, are also regarded as effective ways for interference management [44].

2.3.1 Channel State Information at the Transmitters

Channel state information (CSI) is a set of transmission parameters including the precoding matrix indicator, rank indicator and channel quality indicator, that is reported by UTs corresponding to one or more transmission hypothesises [44]. The acquisition of CSI at the transmitters (CSIT) varies for different scenarios,

Table 2.1: CSI acquisition for different scenarios

Scenarios	**CSI**	**CSI acquisition at transmitters**
TDD	Local users' CSI CSI of other users in the cluster	Exploiting channel reciprocity Overhearing from reverse links
FDD	Local users' CSI CSI of other users in the cluster	Feedback channel Inter BS communications

as summarized in Table 2.1 [44]. In a time division duplex (TDD) scenario, the individual BSs can not only acquire CSI of its own UTs by directly exploiting the channel reciprocity (i.e., employ uplink channel estimate for downlink transmission), but also estimate its crosstalk channels to UTs in other cells within the coordinating cluster by overhearing from their reverse links. Whereas in a frequency division duplex (FDD) scenario, the channels are firstly estimated by the local UTs and then quantized and fed back to the individual BSs through feedback channel. Whilst, the CSIT acquisition of UTs in other coordinated cells can be accomplished through pilot sending by the corresponding BS and inter-BS information exchange via fronthaul link.

The acquisition of accurate knowledge of CSIT, in terms of either the downlink channel vectors, i.e., instantaneous CSI in slow-fading scenarios, or the downlink channel covariance matrices, i.e., statistical CSI in fast-fading scenarios, is necessary to take advantage of multiple antenna techniques and is essential for BSs to design effective downlink transmission strategies such as power optimization scheme. Numerous downlink beamforming designs based on the assumption that the CSI can be perfectly acquired by BSs in real-time, have been studied by the research community [4, 11, 12, 46, 47].

The practical rapidly changing wireless environment, however, can result in outdated estimates in both TDD and FDD scenarios and make the provision of perfect CSI extremely difficult, thus only the imperfect CSI can be acquired at BSs as CSIT is contaminated with unknown errors [48]. The CSI obtained by channel reciprocity in the TDD system is free of quantization and feedback compression errors [44], and the CSIT imperfection in a FDD system might be a consequence of estimation

and quantization errors, and could be outdated and affected by erroneous feedback [49]. In the case that the channel estimation is accurate enough while the amount of feedback bits is limited, the quantization error will be the dominant uncertainties. On contrary, when the feedback rate is unlimited but the channel estimation is not accurate or is outdated, the errors will be dominated by estimation errors. Since the downlink beamforming designs based on perfect CSIT may no longer guarantee the QoS requirements at UTs, various robust beamforming designs have been introduced to provide robustness against CSI errors in wireless communication networks.

The CSI imperfections are usually assumed to have certain properties, either in terms of shapes of uncertainty regions, or statistics. Beamforming designs based on deterministic model or probabilistic model are two methodologies of particular interest. The deterministic model assumes that the CSI perturbations are confined or bounded within an uncertainty region to ensure the worst-case robustness [3, 10, 49–57], e.g., $\mathbf{e}_w^H \mathbf{R}_w \mathbf{e}_w \leq 1$, where $\mathbf{R}_w \succ 0$ specifies the shape and size of the ellipsoid. However, the worst-case optimization to guarantee certain performance for all uncertain channels from a specified region may sometimes be conservative in practice. The probabilistic model (also known as chance-constrained or outage-constrained approach) is usually adopted for channel estimation errors [9, 55, 58–63], where the uncertainty region of the CSI errors are modeled to be statistically unbounded according to some known distributions, e.g., $[\mathbf{e}_w]_t \sim \mathbb{CN}(0, \sigma_t^2)$, and σ_t^2 is the error variance.

2.3.2 Coordinated Transmission

In the past few years, coordinated transmission, i.e., multiple transmission points collaboratively serve the individual UTs, has been recognized as a key enabling technique for future wireless communication networks due to its potential benefits of advanced intercell interference mitigation techniques to substantially improve system throughput, in particular for the cell-edge users [64]. The coordinated transmission can be implemented in both FDD and TDD scenarios [44], and can be utilized for

different deployment scenarios, e.g., homogeneous macro networks with inter-site or intra-site collaboration, and heterogeneous networks with low power picocells or remote radio heads within the macrocell coverage area [65]. As shown in Fig. 2.4,

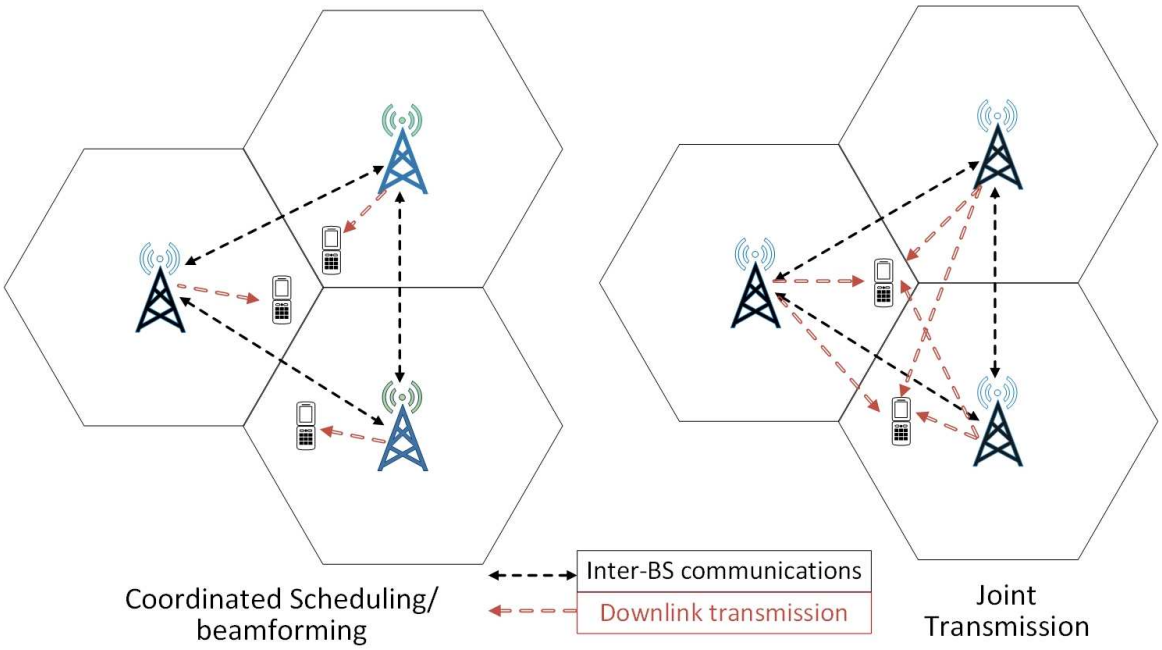


Figure 2.4: Illustration of levels of collaboration amongst BSs.

there are two levels of collaboration amongst BSs in general: joint transmission (JT) with cooperation at signal level, and coordinated scheduling/ beamforming (CS/CB) with coordination at beamforming level [11]. In the JT scenario, multiple fully cooperative BSs act as networked multiple-input multiple-output (MIMO) and adaptively form up clusters to simultaneously transmit data to a single UT with appropriate beamforming weights to improve the received QoS in a time-frequency resource [20]. The JT can be coherently or non-coherently depending upon whether the beamforming from multiple BSs is jointly designed to achieve coherent combining in the wireless channel. This scenario, nevertheless, is more sensitive to the accurate CSI measurements and requires all UTs' data and full CSI reference signals to be circulated among BSs. Sparse beamforming technique for partial cooperation with adaptive BS cooperation clustering, where only part of BSs participate in the JT

to individual UTs on the basis of fronthaul link capacity, is considered as a viable solution to capacity-constrained fronthaul links [6, 47, 66–68]. In contrary, in CS/CB schemes, the decisions of UT scheduling/beamforming are made in a coordinated manner among the cooperating BSs for interference avoidance, while the data for a given UT is only available at and transmitted from one BS. The CS/CB requires rather small information sharing among cooperating BSs and the individual BS only coordinates transmission to its intra-cell UTs using local CSI and a strict CS across cells to alleviate ICI [65]. Although the CS/CB significantly relaxes the fronthaul link capacity via avoidance of UTs' data sharing, it still inflicts a considerable signalling overhead due to its need to full CSI and/or a strict CS to secure the QoS for cell-edge UTs.

A recent emerging trend for implementation of coordinated transmission schemes is to physically detach the baseband processing units from conventional BSs and group them into a centralized cloud computing processor (CP). The functionality of the CP is to execute all scheduling and baseband signal processing, e.g., coordination transmission designs. The remaining remote radio heads that merely perform radio frequency operations, e.g., high frequency signal generation and power amplification, are connected to the CP via finite-capacity low-latency fronthaul links. This promising architecture, known as cloud radio access network (C-RAN), reduces the operating expense and avoids significantly the capital expenditure for hardware upgrade and deployment [69, 70]. However, the resulting immense fronthaul information exchange overhead in such a centralized implementation may be infeasible for practical capacity-constrained fronthaul links [71]. Accordingly, coordinated transmission in a distributed manner, e.g., decentralized coordinated multipoint system and decentralized radio access network, where the individual BSs independently design beamforming vectors based on locally attained information or with limited information exchange among BSs, e.g., sharing only the key intercell coupling parameters, has attracted the attention of researchers in recent years [3, 11, 21].

2.3.3 Benchmark Cooperative Beamforming Designs

The benchmark downlink centralized non-robust beamforming design for BSs cooperated at beamforming level has been presented in Section 2.2.4.

In the sequel, a benchmark worst-case robust coordinated beamforming design proposed in [51] will be introduced. Similar to Section 2.2.4, let us consider N BSs equipped with M antennas cooperated at beamforming level, transmit to K single-antenna UTs over a shared bandwidth. The true channel vector $\mathbf{h}_{ijk}$ from BS i to UT k in cell j can be modeled as $\mathbf{h}_{ijk} = \hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk}\ \forall i,j,k$, where CSI errors are assumed to be bounded within an elliptic uncertainty region, as $\mathbf{e}_{ijk}^H \mathbf{R}_{ijk} \mathbf{e}_{ijk} \leq 1$, and $\mathbf{R}_{ijk} \succ 0$ specifies the shape and size of the ellipsoid. A typical downlink robust power minimization problem can be formulated as

$$\begin{aligned} \min_{\mathbf{w}_{ik},\forall i,k} \quad & \sum_{i=1}^{N}\sum_{k=1}^{K} \|\mathbf{w}_{ik}\|^2 \\ \text{s.t.} \quad & \frac{|\mathbf{h}_{iik}^H \mathbf{w}_{ik}|^2}{\sum_{n\neq k}^{K} |\mathbf{h}_{iik}^H \mathbf{w}_{in}|^2 + \sum_{j\neq i}^{N}\sum_{m=1}^{K} |\mathbf{h}_{jik}^H \mathbf{w}_{jm}|^2 + \sigma_{ik}^2} \geq \gamma_{ik}, \quad \forall i,k, \\ & \mathbf{e}_{ijk}^H \mathbf{R}_{ijk} \mathbf{e}_{ijk} \leq 1, \qquad \forall i,j,k, \end{aligned} \tag{2.20}$$

Introducing slack variables $\{p_{ijk}\}_{i,j,k} \in \mathbb{R}$ and defining the rank-one positive semidefinite matrix $\mathbf{W}_{ik} = \mathbf{w}_{ik}\mathbf{w}_{ik}^H$, the problem in (2.20) can be reformulated as

$$\begin{aligned} \min_{\mathbf{W}_{ik},\forall i,k} \quad & \sum_{i=1}^{N}\sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{W}_{ik}\right) \\ \text{s.t.} \quad & \left(\hat{\mathbf{h}}_{iik} + \mathbf{e}_{iik}\right)^H \left(\gamma_{ik}^{-1}\mathbf{W}_{ik} - \sum_{\substack{n\neq k,\\ n=1}}^{K} \mathbf{W}_{in}\right) \left(\hat{\mathbf{h}}_{iik} + \mathbf{e}_{iik}\right) \geq \sum_{\substack{l\neq i,\\ l=1}}^{N} p_{lik} + \sigma_{ik}^2,\ \forall i,k, \\ & p_{ijk} \geq \left(\hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk}\right)^H \sum_{m=1}^{K} \mathbf{W}_{im} \left(\hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk}\right),\ \forall i, j\neq i, k, \\ & \mathbf{e}_{ijk}^H \mathbf{R}_{ijk} \mathbf{e}_{ijk} \leq 1, \quad \forall i,j,k, \\ & \mathbf{W}_{ik} \succeq 0, \quad \mathrm{rank}\left(\mathbf{W}_{ik}\right) = 1, \quad \forall i,k. \end{aligned} \tag{2.21}$$

By applying the following lemma,

Lemma 2.3.1. *(S-Procedure [8]) The implication* $\mathbf{e}^H\mathbf{A}_1\mathbf{e} + 2\Re(\mathbf{b}_1^H\mathbf{e}) + d_1 \leq 0 \Rightarrow \mathbf{e}^H\mathbf{A}_2\mathbf{e} + 2\Re(\mathbf{b}_2^H\mathbf{e}) + d_2 \leq 0$*, where* $\mathbf{A}_i \in \mathbb{H}^{M\times M}$*,* $\mathbf{b}_i \in \mathbb{C}^M$*,* $d_i \in \mathbb{R}$ *and* $\mathbf{e} \in \mathbb{C}^{M\times 1}$*, holds if and only if there exists a* $\mu \geq 0$ *such that*

$$\begin{bmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^H & d_2 \end{bmatrix} \preceq \mu \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^H & d_1 \end{bmatrix},$$

and relaxing the non-convex rank-one constraints of $\text{rank}(\mathbf{W}_{ik}) = 1$, the problem in (2.21) can be rewritten in SDP form, as

$$\begin{aligned}
\min_{\mathbf{W}_{ik},\boldsymbol{\Phi}_{ik}} \quad & \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i} \text{tr}\left(\mathbf{W}_{ik}\right) \\
\text{s.t.} \quad & \begin{bmatrix} \mu_{ik}\mathbf{R}_{iik} + \boldsymbol{\Phi}_{ik} & \boldsymbol{\Phi}_{ik}\hat{\mathbf{h}}_{iik} \\ (\boldsymbol{\Phi}_{ik}\hat{\mathbf{h}}_{iik})^H & -\mu_{ik} + \hat{\mathbf{h}}_{iik}^H\boldsymbol{\Phi}_{ik}\hat{\mathbf{h}}_{iik} - \sum\limits_{\substack{l\neq i,\\ l=1}}^{N} p_{lik} - \sigma_{ik}^2 \end{bmatrix} \succeq 0, \\
& \mu_{ik} \geq 0,\ \forall i,k, \\
& \begin{bmatrix} \mu_{ijk}\mathbf{R}_{ijk} - \sum\limits_{m=1}^{K}\mathbf{W}_{im} & -\sum\limits_{m=1}^{K}\mathbf{W}_{im}\hat{\mathbf{h}}_{ijk} \\ (-\sum\limits_{m=1}^{K}\mathbf{W}_{im}\hat{\mathbf{h}}_{ijk})^H & -\mu_{ijk} - \hat{\mathbf{h}}_{ijk}^H\sum\limits_{m=1}^{K}\mathbf{W}_{im}\hat{\mathbf{h}}_{ijk} + p_{ijk} \end{bmatrix} \succeq 0, \\
& \mu_{ijk} \geq 0,\ \forall i, j\neq i, k, \\
& \mathbf{W}_{ik} \succeq 0,\ \forall i,k,
\end{aligned} \tag{2.22}$$

where $\boldsymbol{\Phi}_{ik} = \gamma_{ik}^{-1}\mathbf{W}_{ik} - \sum\limits_{\substack{n\neq k,\\ n=1}}^{K}\mathbf{W}_{in}$.

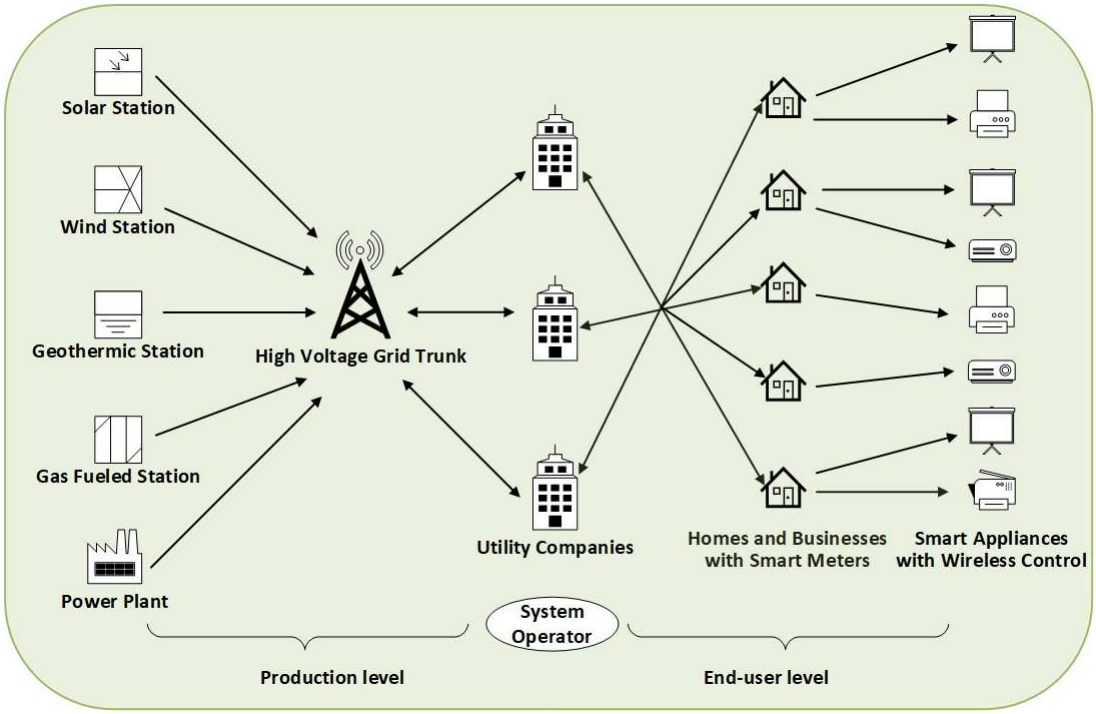


Figure 2.5: A typical smart grid system.

2.4 Energy Trading and Smart Grid

In recent years, the power grid infrastructure has experienced an innovation from the conventional centralized fossil-fuel-based electric grid topology to the highly controllable and distributed smart grid [72]. The smart grid is an electrical grid that includes a variety of operational and energy measures including smart meters, smart appliances, distributed resources and generation, timely information and control options as well as energy efficient and demand-side resources [73]. A typical smart grid system is illustrated in Fig. 2.5, where two-way communication between the utility and its end-users, as well as the sensing along the transmission lines are the key features of the smart grid that provide opportunity to move the energy industry into a new era of reliability, availability, and efficiency. The benefits associated with the smart grid can be concluded as follows:

- More efficient transmission of electricity and quicker restoration of electricity after power disturbances

- Increased integration of large-scale renewable energy systems and better integration of end-user-owned energy generation systems
- Reduced operations and management costs for utilities, and ultimately lower energy costs for end-users
- Better load balancing and peak curtailment via dynamic pricing
- Greater flexibility in network topology and operational strategies for both the suppliers and the end-users
- Improved security and reliability

With the advancement of smart gird technology, two-way energy and timely information flows become viable between the distributed loads, e.g, BSs, and the grid. Energy trading with the grid is gradually becoming a profit making option for both the suppliers and the end-users, and the sophisticate and flexibility in operational strategies provide opportunities for enabling more energy-efficient power networks [74]. As a specific type of distributed loads, BSs in wireless communications networks can accordingly be implemented with the two-way energy trading with the grid to more efficiently utilize their locally generated renewable energy such that the energy cost in a long run can be further reduced [73, 75, 76]. The solutions to the wireless channel random dynamism, the intermittent nature of renewable energy supply and the significant electricity price fluctuation in the wireless communications networks powered by hybrid smart grid and renewable energy, are currently of great interest for the researchers [7, 26–28, 72, 77–79].

2.5 Reinforcement Learning and Multi-armed Bandit Problem

Reinforcement learning is a subfield of machine learning that is typically formulated as a Markov decision process (MDP) comprising of an agent, a set of

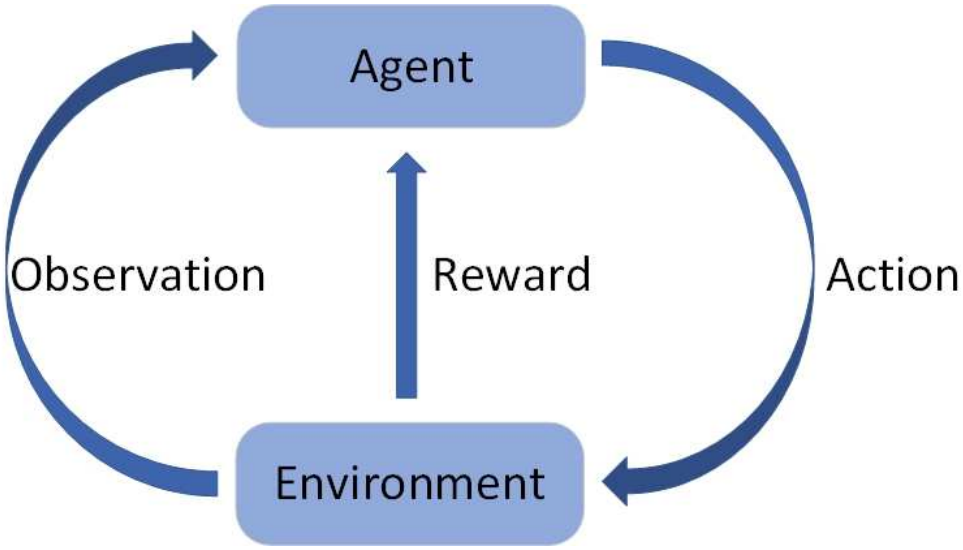


Figure 2.6: A typical frame of a reinforcement learning scenario.

environment and agent states, and a set of actions of the agent, as shown in Fig 2.6. The algorithm attempts to find a policy that maps the environment to the actions the agents take in the environment, with the objective of maximizing the long-term cumulative reward [80]. The reinforcement learning has numerous applications in many disciplines such as control theory, game theory, optimization, multi-agent system and telecommunications. Unlike supervised learning that presents the correct input and output pairs and explicitly corrects the sub-optimal actions, reinforcement learning focuses on online performance that finds a trade-off between exploration of the uncharted territory and exploitation of current knowledge. The agent, in general, is expected not only to take into account the immediate reward, but also to evaluate the consequences of its actions on the future in order to maximize its long-term performance [81]. The trade-off between exploration and exploitation in reinforcement learning has been thoroughly studied through the multi-armed bandit (MAB) problem and in finite MDP [82].

2.5.1 Q-learning for Markov Decision Process

Q-learning is one of the most widely used reinforcement learning techniques that can handle problems with stochastic transitions and rewards, and can eventually find an optimal action-selection policy for any given finite MDP [80]. As introduced above, a MDP involves an agent, a set of states S and a set of actions per state A. The objective of the agent is to maximize its accumulative reward via value iteration update of weighted sum of the expected values. An action-value function, i.e., Q-function, $Q : S \times A \to \mathbb{R}$, is the expected return for a state-action pair for a given policy [81]. At each time t, the agent selects an action $a_t \in A$ in a specific state s_t, observes a reward r_t, transits to a new state s_{t+1} based on s_t and a_t, and update the action-value Q-function, as

$$Q_{t+1}(s_t, a_t) \leftarrow (1-\alpha) \cdot \underbrace{Q_t(s_t, a_t)}_{\text{old value}} + \underbrace{\alpha}_{\text{learning rate}} \overbrace{\left(\underbrace{r_t}_{\text{reward}} + \underbrace{\gamma}_{\text{discount factor}} \underbrace{\max_{a' \in A}(Q_{t+1}(s_{t+1}, a'))}_{\text{estimated future value}} \right)}^{\text{learned value}},$$

until a final or terminate state is achieved. Once the optimal action-value function Q^* is estimated, the agent can select the optimal actions by using a greedy policy [81].

2.5.2 Multi-armed Bandit Problem

The MAB problem is formally equivalent to a one-state MDP. MAB problem is a class of sequential decision making problems and has been extensively studied in probability theory and machine learning. In practice, MAB has been used to model the problems of adaptive routing and server selection in networks, and the click-through probabilities optimization for online advertising. The classical MAB problem is formulated as a system of J arms (or actions), each having an unknown probability distribution of the reward with an unknown mean specific to that arm [83]. The agent iteratively plays one arm per round and observes the associated

reward. The task is to repeatedly play these arms in multiple rounds such that the sum of rewards is as close to the reward of always playing the optimal arm as possible [84]. In each round, the player may lose some reward (also known as regret) due to the selection of the played arm rather than the best arm. The MAB problem requires a trade-off between attempting new arms to further increase knowledge, known as exploration, and selecting the best-possible arm so far based on the knowledge already acquired, known as exploitation [85].

The MAB algorithm is thereby to iteratively optimize the decisions among a set of arms, i.e., decide which arm to play, how many times to play each arm and in which order to play them, in a sequence of rounds so that its accumulated regret over the time horizon is minimized, or its long-term accumulative reward is maximized. An algorithm is said to solve the MAB problem if the resulting regret at the n-th round can match the lower bound of $\text{Regret}_n = O(\log n)$. Several strategies that provide an approximate solution to the bandit problem are briefly introduced as follows

- Epsilon-greedy strategy: The best-possible arm (based on previous observations) is always played except for a proportion ϵ, when an arm is selected uniformly at random.
  - Epsilon-decreasing strategy: Value of ϵ decreases with time, resulting in highly explorative behaviour at beginning and increased exploitative behaviour as learning progresses.
  - Adaptive epsilon-greedy strategy based on value differences [82]: Instead of manual tuning, the value of ϵ is adaptive to the learning progress and environment variations.
- Pricing strategies: A price, e.g., the sum of the expected reward plus an estimation of extra future rewards, is established for each arm and the arm with highest price is always played.

Upper confidence bound (UCB) algorithm, as presented in Algorithm 2.5.1, is one of the most commonly used algorithms to solve MAB problem that automatically

trades off between exploration and exploitation via the perturbation term of $\sqrt{\frac{2\ln n}{n_i}}$.

Algorithm 2.5.1. *Upper Confidence Bound Algorithm*

1: **Initialize***: Play each arm once.*

2: **Loop**

3: *Play the i-th arm that maximizes* $\bar{x}_i + \sqrt{\frac{2lnn}{n_i}}$,

where $\bar{x}_i$ *is the average mean reward of arm* i,

n_i *is the number of times arm* i *has been played so far,*

n *is the total number of plays done so far.*

The expected regret of UCB algorithm after n plays is at most [84]

$$[8 \sum_{i:\mu_i \leq \mu^*} (\frac{\ln n}{\mu^* - \mu_i})] + (1 + \frac{\pi^2}{3})(\sum_{i=1}^{J} \mu^* - \mu_i), \tag{2.23}$$

where J is the total number of arms, μ_i is the reward expectation for arm i and μ^* is the maximal element. UCB algorithm achieves $O(\log n)$ regret without any preliminary knowledge about the reward distributions.

2.5.3 Variations of MAB problem

The variations of MAB problem can be briefly summarized as follows [80]

- Binary or Bernoulli MAB problem: A reward of one is issued with probability p, and a reward of zero otherwise.
- Restless bandit problem: Each arm represents an independent Markov machine and each time an arm is played, the state of that played arm (and even non-played arms) evolve over time.
- Adversarial bandit: An agent chooses an arm at each iteration, and an adversary simultaneously chooses the payoff structure for each arm.

Another variant of the MAB problem is the combinatorial multi-armed bandit (CMAB) problem, which can be considered as a combinatorial generation of classic

MAB problem. The CMAB problem is defined as a system consists of J base arms, whose outcomes follow certain unknown joint distribution, where a set of N base arms (also known as a super arm), $N \subset J$, is played simultaneously and the reward of each arm is observed individually at each round. The objective is to maximize the long-term accumulated reward via a trade-off between exploring new super arms that might yield a better reward and exploiting the best-possible super arm that is associated with the highest reward so far based on knowledge acquired from the previous rounds [85]. The combinatorial UCB algorithm, which is an extension to the UCB algorithm for the classical MAB problem, is used to compute the optimal super arm with respect to the input and is proved to have the regret being bounded by $O(\log n)$ after n rounds [84].

The applications of bandit model to wireless communications networks have been recently studied in the literature [12, 85–89]. Due to the combinatorial nature of distributed energy transmission from the grid to the BSs, the price-aware energy management problem studied in this thesis is classified as a CMAB problem.

2.6 Related Works

The authors in [6] employ sparse beamforming technique and propose a non-robust user-centric clustering beamforming design to account for the fronthaul capacity constraints in a centralized downlink C-RAN. The authors in [5] combine the classic energy efficient coordinated beamforming design with SWIPT technique and develop a centralized non-robust transmission strategy based on sparse beamforming technique in downlink C-RAN. In [90], an energy-efficient resource allocation approach based on cross-tier interference reduction is introduced for two-tier macrocell/femtocell networks. In [11], the authors propose a decentralized iterative algorithm using subgradient method for non-robust sum-power minimization and max-min SINR beamforming design via limited inter-BS communications in multicast multi-cell networks, based on the assumption of perfect knowledge of CSIT. The authors in [91] investigate a max-min weighted SINR problem under weighted sum

power constraint. The problem is solved based on uplink-downlink duality in a distributed manner that only requires statistical information. Another distributed approach to a sum-power minimization problem in a coordinated network is proposed in [92] where an iterative algorithm is introduced to jointly minimize a linear combination of total transmit power and weighted ICI using statistical CSI. However, the designs take no consideration of any channel uncertainties, which can no longer guarantee the QoS constraints at UTs and may result in unpredictable results in practice.

Assuming that the uncertainty region of CSI perturbations is bounded, the authors in [3, 10, 49, 50] investigate the robust sum-power minimization problem subject to worst-case QoS constraints at UTs in a distributed manner in downlink multi-cell coordinated networks. Based on [11], the authors in [3] and [50] introduce distributed subgradient iterative algorithms for downlink sum-power minimization transmission designs via limited signaling among BSs and provide worst-case robustness against CSI imperfections. In [10], the authors propose a distributed algorithm based on the principle of alternating direction method of multipliers technique to minimize the weighted sum power with limited fronthaul information exchange between BSs in multi-cell network under the assumption of hyper-sphere bounded CSI errors. Although the robust design on the basis of deterministic model guarantees the worst-case robustness against CSI uncertainties, it is conservative and may require higher transmit power to count for the worst-case QoS. This is due to the fact that the worst-case is a rare occurrence in practice and the realistic wireless network can tolerate occasional QoS outages.

On the contrary, the stochastic model of CSI imperfection provides a less conservative solution in terms of energy efficiency. The CSI uncertainties are modeled to be statistically unbounded with some known distribution and the robust design based on stochastic model satisfies the QoS requirements at UTs with a certain probability. [58] investigates a beamforming design to jointly coordinate the aggregated transmit power and overall ICI pricing with an outage probability

threshold being assigned to each SINR constraint. The design provides robustness against the second order statistical CSI errors and the authors assume that the statistical average of total ICI can be accurately estimated by the UTs and then updated to the local BS. Their designs, nevertheless, take no account of guaranteeing the transmit power to be within the available power budget at individual BSs, which may lead to an infeasible solution in a realistic scenario. [93] studies a distributed outage probability based rate utility maximization problem under individual BSs' power constraints with a limited amount of information exchange among BSs. Assuming instantaneous CSI errors are Gaussian distributed and employing the Bernstein-type inequality method, the authors in [60] introduce an outage probability based robust transmission design for overall power minimization problem in a single cell scenario. Another Bernstein-type inequality method based robust transmission design for instantaneous CSI is proposed in [9], where the total transmit power is minimized and the QoS constraints for UTs are satisfied above a certain outage probability threshold. However, the Bernstein-type inequality method obtains feasible worst-case solutions by approximating the probabilistic constraints with their convex upper bounds, which is conservative for practical scenarios.

In addition to energy efficiency and ICI management for green communications, powering BSs with renewable energy generation and smart grid to compensate for the wireless network dynamics and electricity price fluctuation, has attracted the attention of researchers recently [7, 12, 27, 28]. Assuming the availability of hourly varying profiles of BSs' energy demand and renewable generation as well as the day-ahead knowledge of hourly-varying electricity prices, [77] minimizes the electricity bill at BSs powered jointly by smart grid and locally harvested solar energy. In [26, 94], two-way energy trading between the BSs and the grid in a coordinated multipoint (CoMP) system is studied based on convex optimization techniques and concluded that the joint management of energy trading by fully cooperative BSs reduces the total energy cost. Partial cooperation based on sparse beamforming technique is proposed in [4, 46] to account for limited-capacity fronthaul links connecting the

CP and BSs in CoMP systems, whilst two-way energy trading with the grid is performed. Without the involvement of online learning concept, the authors in [78] study energy trading amongst a set of storage units and the grid from the perspective of noncooperative game theory and propose an algorithm that achieves at least one Nash equilibrium point. The authors in [72] formulate the system as a simplified two-level Stackelberg game in the smart-grid-powered green CoMP system. In [28], the authors conclude the demand-side power management solutions for a single BS in the smart-grid-powered green CoMP powered by hybrid renewable energy and electrical grid, whereas the scenario of multiple BSs is not considered therein. The authors in [95] study energy allocation problem for renewable energy powered BSs using a noncooperative game powered by hybrid renewable energy and electrical grid. However, their designs require statistics of the system dynamics, which is not a realistic assumption in practice. Furthermore, none of these designs considered the impact of online learning on cost-aware proactive energy management, or provided adaption to the wireless system dynamics without requiring upfront statistical knowledge. Requiring no prior knowledge of traffic, the authors in [96] develop an adaptive resource management in vehicular access network. Assuming prior knowledge of statistical distribution of upcoming energy price and demand load, the authors in [7] propose an online learning algorithm for stochastic storage management in smart grid rather than cellular network based on MDP model. Using stochastic dual-subgradient method based optimization rather than online learning over infinite time horizon, the authors in [27] propose a dynamic energy management scheme for the smart-grid-powered CoMP, where BSs are fully cooperated and governed by a CP. The authors in [12] first introduce the application of CMAB as an online learning approach to energy management design based on sparse beamforming technique in a simplified network scenario, where randomness of the renewable energy generation and wireless channel dynamics are relaxed, and the exploration is in single direction. Furthermore, their proposed design takes no consideration of the long-term effect or the deployment of energy storage devices, and a full exploration CMAB

algorithm without an efficient trade-off strategy between the exploration and the exploitation is proposed.

2.7 Concluding Remarks

This chapter provides some mathematical preliminaries such as convex optimization as well as a general overview of cooperative transmission in downlink multi-cell interference networks, which will be adopted in the subsequent chapters. Furthermore, two methodologies of CSI imperfection modelling, i.e., probabilistic model and deterministic model, that will be employed in Chapter 3 and Chapter 4, respectively, are introduced in Section 2.3.1, followed by the introduction of two benchmark cooperative beamforming designs in Section 2.2.4 and Section 2.3.3, respectively. Section 2.4 and Section 2.5 present recent advances in energy trading and smart grid as well as reinforcement learning and the bandit problem, respectively, which will be applied in cost-aware energy management designs in Chapter 5 and Chapter 6. Finally, Section 2.6 presents a literature review of the state-of-the-art energy management strategies for green communications from both energy efficient and cost efficient perspectives.

Chapter 3

Robust Outage Probability based Distributed Beamforming for Multicell Interference Networks

3.1 Introduction

This chapter investigates energy-efficient transmission strategies with interference control for green communications in multi-cell interference networks. Taking the capacity-limited fronthaul links in practical scenario into account, two robust coordinated transmission strategies are proposed to minimize the aggregate downlink transmit power in a distributed manner in the presence of imperfect channel state information (CSI). Due to the fact that worst-case is a rare occurrence in practical network, the problems are constrained to satisfying a set of signal-to-interference-plus-noise-ratio (SINR) requirements at individual user terminals (UTs) with certain predefined SINR outage probabilities. The proposed strategies provide robustness against the second order statistical CSI estimation errors and the instantaneous CSI uncertainties, respectively.

3.1.1 Main Contributions

Both of the problems are numerically intractable due to the cross-link coupling effect across multiple base stations (BSs) operating under the same frequency bandwidth and the robust SINR constraints that involve the second order statistical or instantaneous CSI estimation errors, respectively. By employing Schur complement, S-procedure, cumulative distribution function (CDF) of standard normal distribution and semidefinite relaxation (SDR) technique, the intractable problems are first converted to the tractable semidefinite programming (SDP) form

with linear matrix inequality (LMI) constraints that can be solved in a centralized fashion. Then, an iterative subgradient based learning algorithm is introduced to decompose the multicell-wise general problem into a set of independent equivalent parallel subproblems at individual BSs. The subgradient based learning algorithm allows the individual BSs to exchange some key intercell coupling parameters and gradually learn the intercell interference (ICI) imposed from other BSs, such that the ICI coordination among BSs can be achieved with a light inter-BS communications overhead.

Simulation results reveal that the proposed outage probability based transmission strategies outperform the distributed worst-case bounded error designs in [10] and [50], and an outage probability based robust beamforming design in [9] for most cases in terms of providing better energy efficiency and expending SINR operational range.

3.1.2 Organization

The rest of this chapter is organized as follows. Section 3.2 introduces an outage probability based distributed robust transmission strategy that provides robustness against the second order statistical CSI uncertainties with a set of outage levels. In Section 3.2.1, the system model and optimization problem formulation will be introduced. Then, the original numerically intractable problem is transformed into a SDP form with LMI constraints in 3.2.2 and decomposed via inter BS learning iterations in section 3.2.3, followed by fronthaul signalling load and computational complexity analysis. Section 3.3 introduces a distributed robust transmission strategy with SINR certain outage probabilities against instantaneous CSI estimation error. Section 3.3.1 introduces the system model and sum-power minimization problem formulation. In section 3.3.2, the original problem is first reformulated as a probabilistic constrained optimization problem and then transformed into SDP form with LMI constraints. Then, the general problem is decomposed and solved via projected subgradient learning method in section 3.3.3. The simulation results are

analyzed in Section 3.4. Finally, Section 3.5 concludes the chapter.

3.2 Robust Transmission in Multicell Networks with Probabilistic Constraints involving Statistical CSI Uncertainties

3.2.1 System Model and Problem Formulation

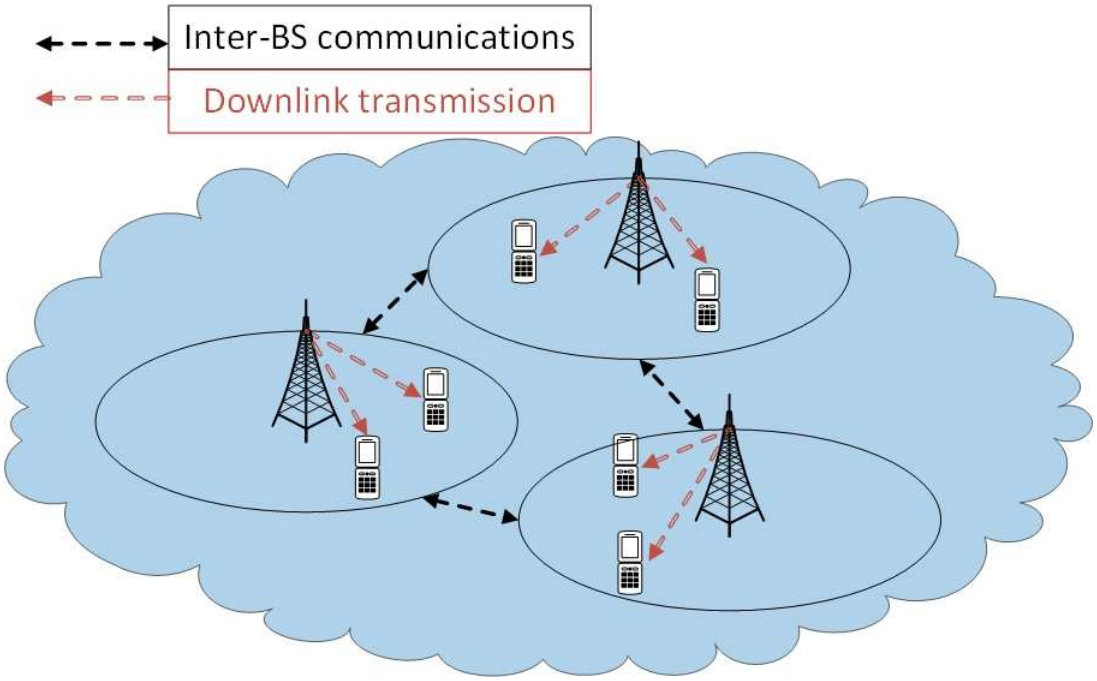


Figure 3.1: Illustration of system scenario.

Consider a downlink multicell network with a coordinated cluster of N cells, as shown in Fig. 3.1. Each cell consists of a BS equipped with an array of M antenna elements transmitting to K single-antenna UTs over a shared bandwidth. Let the set of indexes for the BSs and the UTs be denoted as $\mathcal{L}_b = \{1, \cdots, N\}$ and $\mathcal{L}_i = \{1, \cdots, K\}$, respectively,. Let BS_i, $i \in \mathcal{L}_b$ indicate the BS in the i-th cell, and UT_{ik}, $k \in \mathcal{L}_i$ represent the k-th UT in the i-th cell. Then, the signal received by

UT_{ik} is given by

$$z_{ik} = \mathbf{h}_{iik}^H \mathbf{w}_{ik} s_{ik} + \sum_{\substack{n \neq k, \\ n \in \mathcal{L}_i}} \mathbf{h}_{iik}^H \mathbf{w}_{in} s_{in} + \sum_{\substack{j \neq i, \\ j \in \mathcal{L}_b}} \sum_{m \in \mathcal{L}_i} \mathbf{h}_{jik}^H \mathbf{w}_{jm} s_{jm} + n_{ik}, \tag{3.1}$$

where s_{ik} indicates the data symbol for UT_{ik}, $\mathbf{w}_{ik} \in \mathbb{C}^{M \times 1}$ and $\mathbf{h}_{ijk} \in \mathbb{C}^{M \times 1}$ denote the beamforming vector for UT_{ik} and the channel vector from BS_i to UT_{jk}, respectively. Note that the terms in the right hand side of (3.1), respectively, represent the desired signal, the total intra-cell interference, the aggregate ICI and the zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise at UT_{ik}, e.g., $n_{ik} \sim \mathbb{CN}(0, \sigma_{ik}^2)$.

Let $\hat{\mathbf{C}}_{ijk} = \mathbb{E}(\hat{\mathbf{h}}_{ijk} \hat{\mathbf{h}}_{ijk}^H) \in \mathbb{C}^{M \times M}$ denote the estimated channel covariance matrix of UT_{jk}, as seen by the i-th BS. Also let $\mathbf{\Delta}_{ijk} \in \mathbb{C}^{M \times M}$ represent the corresponding error matrix, where the (c, d)-th entry of $\mathbf{\Delta}_{ijk}$ is distributed as $[\mathbf{\Delta}_{ijk}]_{cd} \sim \mathbb{CN}(0, \sigma_{cd}^2)$. Then, the true channel covariance matrix $\mathbf{C}_{ijk}$ can be modeled as

$$\mathbf{C}_{ijk} = \hat{\mathbf{C}}_{ijk} + \mathbf{\Delta}_{ijk}, \qquad \forall i, j, k. \tag{3.2}$$

Assuming the normalized energy of transmitted symbols, i.e., $\mathbb{E}(|s_{ik}|^2) = 1$, and the identical σ_{ik}^2 at all UTs, the SINR at UT_{ik} can be formulated as

$$\mathrm{SINR}_{ik} = \frac{\mathbf{w}_{ik}^H \mathbf{C}_{iik} \mathbf{w}_{ik}}{\sum\limits_{\substack{n \neq k, \\ n \in \mathcal{L}_i}} \mathbf{w}_{in}^H \mathbf{C}_{iik} \mathbf{w}_{in} + \sum\limits_{\substack{j \neq i, \\ j \in \mathcal{L}_b}} \sum\limits_{m \in \mathcal{L}_i} \mathbf{w}_{jm}^H \mathbf{C}_{jik} \mathbf{w}_{jm} + \sigma_{ik}^2}. \tag{3.3}$$

Let us consider a robust problem of minimizing the total transmit power in multi-cell networks under the constraints of satisfying the SINR requirements at individual UTs with certain SINR outage probabilities in the presence of channel estimation errors, as

$$\begin{aligned} \min_{\mathbf{w}_{ik}, \forall i,k} \quad & \sum_{i \in \mathcal{L}_b} \sum_{k \in \mathcal{L}_i} \|\mathbf{w}_{ik}\|^2 \\ \text{s.t.} \quad & \Pr\ (\mathrm{SINR}_{ik} \geq \gamma_{ik}) \geq 1 - \rho_{ik}, \quad \forall i \in \mathcal{L}_b, k \in \mathcal{L}_i, \end{aligned} \tag{3.4}$$

where γ_{ik} is the target SINR requested by UT_{ik}, $\rho_{ik} \in (0,1)$ is the maximum SINR outage probability and $1-\rho_{ik}$ indicates that the individual UTs is guaranteed to achieve its target SINR with probability of $1-\rho_{ik}$ at the least.

In order to account for the coupling effects among the multiple cells, let us begin by introducing slack variables $\{p_{ijk}\}_{i,j,k} \in \mathbb{R}$ that indicates the ICI from BS_i to UT_{jk}, and reformulating the problem in (3.4) as

$$\min_{\mathbf{w}_{ik}, p_{ijk}} \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i} \|\mathbf{w}_{ik}\|^2 \tag{3.5}$$

$$\text{s.t.} \quad \Pr\left(\frac{\mathbf{w}_{ik}^H(\hat{\mathbf{C}}_{iik}+\boldsymbol{\Delta}_{iik})\mathbf{w}_{ik}}{\sum\limits_{\substack{n\neq k,\\ n\in\mathcal{L}_i}} \mathbf{w}_{in}^H(\hat{\mathbf{C}}_{iik}+\boldsymbol{\Delta}_{iik})\mathbf{w}_{in} + \sum\limits_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} + \sigma_{ik}^2} \geq \gamma_{ik} \right) \geq 1-\rho_{ik}, \forall i\in\mathcal{L}_b, k\in\mathcal{L}_i,$$

$$\Pr\left(\sum_{m\in\mathcal{L}_i} \mathbf{w}_{im}^H(\hat{\mathbf{C}}_{ijk}+\boldsymbol{\Delta}_{ijk})\mathbf{w}_{im} \leq p_{ijk} \right) \geq 1-\rho_{ik},\ \forall i\in\mathcal{L}_b, j\neq i, k\in\mathcal{L}_i.$$

3.2.2 Optimization of Problem in (3.5)

The problem in (3.5) is numerical intractable since the inclusion of the second order statistical CSI uncertainties in probabilistic constraints naturally lead to an infinite number of convex sets. Therefore, following the similar principles as in [58], the probabilistic constraints of the problems in (3.5) can be equivalently transformed into a tractable form through the following Lemma.

Lemma 3.2.1. *Let $\boldsymbol{\Delta} \in \mathbb{C}^{M\times M}$ be a Hermitian random matrix with each ZMCSCG element being characterized as $[\boldsymbol{\Delta}]_{cd} \sim \mathbb{CN}(0,\sigma_{cd}^2)$. Then, for any Hermitian matrix $\mathbf{L}$, $\mathbf{L}\in\mathbb{C}^{M\times M}$,*

$$tr(\mathbf{L}\boldsymbol{\Delta}) \sim \mathbb{N}(0, \|\mathcal{D}_{\boldsymbol{\Delta}} vec(\mathbf{L})\|^2),$$
$$tr(\mathbf{L}\boldsymbol{\Delta}) = \|\mathcal{D}_{\boldsymbol{\Delta}} vec(\mathbf{L})\| U,\ U\sim\mathbb{N}(0,1),$$

where $\mathcal{D}_{\boldsymbol{\Delta}} = diag(vec(\boldsymbol{\Sigma}_{\boldsymbol{\Delta}}{}^H))$ and $\boldsymbol{\Sigma}_{\boldsymbol{\Delta}}$ denotes a real-valued $M\times M$ matrix with each

entry $[\mathbf{\Sigma}_{\mathbf{\Delta}}]_{cd} = \sigma_{cd}$.

Proof: Please refer to Appendix A [58].

Let the rank-one positive semidefinite matrix be defined as $\mathbf{W}_{ik} = \mathbf{w}_{ik}\mathbf{w}_{ik}^H$. Also let the first and the second set of constraints in (3.5) be rewritten as follows

$$\Pr\left(\mathrm{tr}(-\mathbf{B}_{ik}\mathbf{\Delta}_{iik}) \leq \mathrm{tr}(\mathbf{B}_{ik}\hat{\mathbf{C}}_{iik}) - \sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} - \sigma_{ik}^2\right) \geq 1-\rho_{ik}, \tag{3.6}$$

$$\Pr\left(\mathrm{tr}(\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im}\mathbf{\Delta}_{ijk}) \leq p_{ijk} - \mathrm{tr}(\hat{\mathbf{C}}_{ijk}\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})\right) \geq 1-\rho_{ik}, \tag{3.7}$$

where $\mathbf{B}_{ik} = \gamma_{ik}^{-1}\mathbf{W}_{ik} - \sum_{\substack{n\neq k,\\ n\in\mathcal{L}_i}}\mathbf{W}_{in}$. By applying Lemma 3.2.1 and the CDF of a standard normal distribution, i.e., $\phi(u) = \Pr(U \leq u) = \frac{1}{2}[1+\mathrm{erf}(\frac{u}{\sqrt{2}})]$, where $U \sim \mathbb{N}(0,1)$, the first and the second constraints in (3.6) and (3.7), respectively, can be expressed as follows

$$\Pr\left(\mathrm{tr}(-\mathbf{B}_{ik}\mathbf{\Delta}_{iik}) \leq \mathrm{tr}(\mathbf{B}_{ik}\hat{\mathbf{C}}_{iik}) - \sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} - \sigma_{ik}^2\right) \tag{3.8}$$

$$= \Pr\left(U \leq \frac{\mathrm{tr}(\mathbf{B}_{ik}\hat{\mathbf{C}}_{iik}) - \sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} - \sigma_{ik}^2}{\|\mathcal{D}_{\mathbf{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik})\|}\right)$$

$$= \frac{1}{2}[1+\mathrm{erf}\left(\frac{\mathrm{tr}(\mathbf{B}_{ik}\hat{\mathbf{C}}_{iik}) - \sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} - \sigma_{ik}^2}{\sqrt{2}\|\mathcal{D}_{\mathbf{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik})\|}\right)] \geq 1-\rho_{ik},$$

$$\Pr\left(\mathrm{tr}(\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im}\mathbf{\Delta}_{ijk}) \leq p_{ijk} - \mathrm{tr}(\hat{\mathbf{C}}_{ijk}\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})\right) \tag{3.9}$$

$$= \Pr\left(U \leq \frac{p_{ijk} - \mathrm{tr}(\hat{\mathbf{C}}_{ijk}\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})}{\|\mathcal{D}_{\mathbf{\Delta}_{ijk}}\mathrm{vec}(\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})\|}\right)$$

$$= \frac{1}{2}[1+\mathrm{erf}\left(\frac{p_{ijk} - \mathrm{tr}(\hat{\mathbf{C}}_{ijk}\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})}{\sqrt{2}\|\mathcal{D}_{\mathbf{\Delta}_{ijk}}\mathrm{vec}(\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})\|}\right)] \geq 1-\rho_{ik},$$

which are equivalent to the following expressions, respectively, as

$$\Theta \geq \sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})\|\mathcal{D}_{\mathbf{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik})\|, \tag{3.10}$$

$$\Upsilon \geq \sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})\|\mathcal{D}_{\mathbf{\Delta}_{ijk}}\mathrm{vec}(\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})\|, \tag{3.11}$$

where $\Theta = \mathrm{tr}(\mathbf{B}_{ik}\hat{\mathbf{C}}_{iik}) - \sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} - \sigma_{ik}^2$ and $\Upsilon = p_{ijk} - \mathrm{tr}(\hat{\mathbf{C}}_{ijk}\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})$.

Lemma 3.2.2. *(Schur Complements [97]) The following second order cone constraint on x*

$$\|Ax+b\| \leq e^T x + d$$

is equivalent to the following LMI form

$$\begin{bmatrix} (e^T x+d)\mathbf{I} & Ax+b \\ (Ax+b)^T & e^T x+d \end{bmatrix} \succeq 0.$$

Applying Lemma 3.2.2 to (3.10) and (3.11), the problem in (3.5) can be reformulated as a SDP form with LMI constraints after relaxing the rank-one constraints of rank $(\mathbf{W}_{ik}) = 1, \forall i \in \mathcal{L}_i, n \in \mathcal{L}_b$, via SDR [38], as

$$\begin{aligned} \min_{\mathbf{W}_{ik},\mathbf{p}_i,\Theta,\Upsilon} \ & \sum_{i\in\mathcal{L}_b} f_i(\mathbf{W}_{ik},\mathbf{p}_i) \triangleq \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i}\mathrm{tr}(\mathbf{W}_{ik}) \\ \text{s.t.} \quad & \begin{bmatrix} \frac{\Theta}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik}) \\ \mathrm{vec}^H(-\mathbf{B}_{ik})\mathcal{D}_{\mathbf{\Delta}_{iik}} & \frac{\Theta}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix} \succeq 0, \\ & \begin{bmatrix} \frac{\Upsilon}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{ijk}}\mathrm{vec}(\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im}) \\ \mathrm{vec}^H(\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})\mathcal{D}_{\mathbf{\Delta}_{ijk}} & \frac{\Upsilon}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix} \succeq 0, \\ & \mathbf{W}_{ik} \succeq 0, \ \forall i \in \mathcal{L}_i, n \in \mathcal{L}_b, \end{aligned} \tag{3.12}$$

where $\mathbf{p}_i \in \mathbb{R}^{NK\times 1}$, $\forall i, j \neq i$, is a real-valued vector that contains the local intercell

coupling variables at the i-th BS, i.e.,

$$\mathbf{p}_i = \Big[\sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{li1}, \sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{li2}, ..., \sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{liK}\Big| p_{ij1}, p_{ij2}, ..., p_{iNK}\Big]^T, \tag{3.13}$$

and the function $f_i(\mathbf{W}_{ik}, \mathbf{p}_i) = \sum_{k\in\mathcal{L}_i} \mathrm{tr}(\mathbf{W}_{ik})$ in (3.12) indicates the dependence of f_i on $\mathbf{p}_i$. The primal problem in (3.12) can now be solved in a centralized fashion. Due to the fact that in a practical scenario, the fronthaul link has limited capacity, in the next section, the problem in (3.12) will be decomposed via primal decomposition [98] to further relax the fronthaul links.

3.2.3 Distributed Optimization of Problem in (3.12)

Let the global intercell coupling variables $\mathbf{p} \in \mathbb{R}^{(N(N-1)+1)K\times 1}$ be defined as

$$\mathbf{p} = \Big[p_{121}, p_{122}, ..., p_{12K}, ..., p_{N11}, ..., p_{NN-1K}\Big|\mathbf{0}_{K\times 1}^T\Big]^T. \tag{3.14}$$

In the sequel, a direction matrix $\mathbf{X}_i$ is introduced to extract $\mathbf{p}_i$ from $\mathbf{p}$, i.e., $\mathbf{p}_i = \mathbf{X}_i\mathbf{p}$, so that the individual BSs can locally design the multicell-wise optimum beams towards its local UTs in a distributed manner. Similar to [3], let us define $\mathbf{X}_i = \begin{bmatrix}\mathbf{A}_i^T & \mathbf{S}_i^T\end{bmatrix}^T \in \{0,1\}^{NK\times(N(N-1)+1)K}$, where $\mathbf{A}_i \in \{0,1\}^{K\times(N(N-1)+1)K}$ and $\mathbf{S}_i \in \{0,1\}^{(N-1)K\times(N(N-1)+1)K}$. The i-th BS constructs $\mathbf{A}_i$ and $\mathbf{S}_i$ by rotating each one of the rows of matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{S}}$, respectively, $(i-1)NK$ and $(i-1)(N-1)K$ times anticlockwise, where

$$\hat{\mathbf{A}} = \Big[\ \mathbf{0}_{K\times(N-1)K}\big|\overbrace{\mathbf{I}_{K\times(N-1)K}}^{(N-1)}\big|\mathbf{0}_{K\times K}\ \Big], \tag{3.15}$$

$$\hat{\mathbf{S}} = \big[\mathbf{I}_{(N-1)K\times(N-1)K}\big|\mathbf{0}_{(N-1)K\times((N-1)^2+1)K}\big]. \tag{3.16}$$

Then the i-th BS extracts the entries of $\mathbf{p}_i$, as

$$\sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} = \mathbf{1}_k^T\mathbf{X}_i\mathbf{p}, \quad \forall k, \tag{3.17}$$

$$p_{ijk} = \mathbf{1}_q^T\mathbf{X}_i\mathbf{p}, \quad \forall j\neq i,k, \tag{3.18}$$

where $q = k + jK$ for $j < i$ and $q = k + (j-1)K$ for $j > i$.

Based on the principle of decomposition theory [98], the primal problem in (3.12) can be decomposed into two levels of optimization, i.e., a lower level at which N subproblems are distributively solved at individual BSs for a fixed global variable $\mathbf{p}$, and a higher level at which a master problem is in charge of updating $\mathbf{p}$. For any fixed global variable $\mathbf{p}$, the equivalent sub-problems at each BS i of problem in (3.12) can be expressed as

$$\begin{aligned}
\min_{\mathbf{W}_{ik}} \quad & f_i(\mathbf{W}_{ik}) = \sum_{k\in\mathcal{L}_i}\mathrm{tr}(\mathbf{W}_{ik}) \\
\text{s.t.} \quad & \mathbf{T}_{ik} = \mathbf{T}'_{ik} - (\mathbf{1}_k^T\mathbf{X}_i\mathbf{p})\mathbf{I}_{(M^2+1)} \succeq 0, \quad \forall k, \\
& \mathbf{T}_{ijk} = \mathbf{T}'_{ijk} + (\mathbf{1}_q^T\mathbf{X}_i\mathbf{p})\mathbf{I}_{(M^2+1)} \succeq 0, \quad \forall j\neq i,k, \\
& \mathbf{W}_{ik} \succeq 0, \quad \forall k,
\end{aligned} \tag{3.19}$$

where

$$\mathbf{T}'_{ik} = \begin{bmatrix} \frac{\mathrm{tr}(\mathbf{B}_{ik}\hat{\mathbf{C}}_{iik})-\sigma_{ik}^2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik}) \\ \mathrm{vec}^H(-\mathbf{B}_{ik})\mathcal{D}_{\mathbf{\Delta}_{iik}} & \frac{\mathrm{tr}(\mathbf{B}_{ik}\hat{\mathbf{C}}_{iik})-\sigma_{ik}^2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix}, \tag{3.20}$$

$$\mathbf{T}'_{ijk} = \begin{bmatrix} \frac{-\mathrm{tr}(\hat{\mathbf{C}}_{ijk}\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{ijk}}\mathrm{vec}(\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im}) \\ \mathrm{vec}^H(\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})\mathcal{D}_{\mathbf{\Delta}_{ijk}} & \frac{-\mathrm{tr}(\hat{\mathbf{C}}_{ijk}\sum_{m\in\mathcal{L}_i}\mathbf{W}_{im})}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix}.$$

Then, the master problem that is in charge of updating the global variable $\mathbf{p}$, is defined as $\min_{\mathbf{p}} \sum_{i\in\mathcal{L}_b} f_i^*(\mathbf{p})$, where f_i^* is the optimal solution to subproblem i in (3.19).

For a fix value of $\mathbf{p}$, the Lagrangian of subproblem i in (3.19) can be expressed as

$$L_i(\{\mathbf{W}_{ik}, \lambda_{ik}\}_k, \{\lambda_{ijk}\}_{j\neq i,k}) = \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\mathbf{W}_{ik}\right) - \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\lambda_{ik}\mathbf{T}_{ik}\right) - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\lambda_{ijk}\mathbf{T}_{ijk}\right), \tag{3.21}$$

where $\lambda_{ik}, \lambda_{ijk} \in \mathbb{H}^{(M^2+1)\times(M^2+1)}$ are the Lagrange multipliers and are positive semidefinite. Since the problem in (3.19) is convex and satisfies the Slaters condition, strong duality holds [8] and the dual function is given by

$$\begin{aligned} \ell_i(\mathbf{p}) \quad &= \inf_{\mathbf{W}_{ik}\succeq 0} L_i = \Xi\left(\{\lambda_{ik}\}_k, \{\lambda_{ijk}\}_{j\neq i,k}\right) \\ &+ \left(\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ik}\mathbf{I})\mathbf{1}_k^T - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ijk}\mathbf{I})\mathbf{1}_q^T\right)\mathbf{X}_i\mathbf{p}, \end{aligned} \tag{3.22}$$

where

$$\begin{aligned} \Xi\left(\{\lambda_{ik}\}_k, \{\lambda_{ijk}\}_{j\neq i,k}\right) = \quad & \inf_{\mathbf{W}_{ik}\succeq 0} \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\mathbf{W}_{ik}\right) - \sum_{k\in\mathcal{L}_i} \mathrm{tr}\big(\lambda_{ik}\mathbf{T}'_{ik}\big) \\ & - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}\big(\lambda_{ijk}\mathbf{T}'_{ijk}\big). \end{aligned} \tag{3.23}$$

Then we can write

$$f_i^*(\mathbf{W}_{ik}^*, \mathbf{p}_i) = f_i^*(\mathbf{p}) = \ell_i^*(\mathbf{p}) = \mathbf{g}_i\mathbf{p} + \Xi\left(\{\lambda_{ik}^*\}_k, \{\lambda_{ijk}^*\}_{j\neq i,k}\right), \tag{3.24}$$

where

$$\mathbf{g}_i = \left(\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ik}^*\mathbf{I})\mathbf{1}_k^T - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ijk}^*\mathbf{I})\mathbf{1}_q^T\right)\mathbf{X}_i. \tag{3.25}$$

It can be easily concluded from (3.24) that for any given $\hat{\mathbf{p}}$,

$$\ell_i^*(\hat{\mathbf{p}}) \geq \ell_i^*(\mathbf{p}) + \mathbf{g}_i(\hat{\mathbf{p}} - \mathbf{p}). \tag{3.26}$$

Therefore, $\mathbf{g}_i \in \mathbb{R}^{1\times(N(N-1)+1)K}$ is the subgradient vector of $\ell_i^*(\mathbf{p})$ and $f_i^*(\mathbf{p}_i)$ obtained for the i-th subproblem. Following the similar steps of analysis as for subproblem i in (3.19), one can easily calculate the global subgradient of $\sum_{i\in\mathcal{L}_b} f_i^*(\mathbf{p}_i)$ obtained for the general problem in (3.12) at a given value of $\mathbf{p}$, as

$$\begin{aligned}\mathbf{g} &= \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ik}^*\mathbf{I})\mathbf{1}_k^T\mathbf{X}_i - \sum_{i\in\mathcal{L}_b}\sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}}\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ijk}^*\mathbf{I})\mathbf{1}_q^T\mathbf{X}_i \\ &= \sum_{i\in\mathcal{L}_b}\left(\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ik}^*\mathbf{I})\mathbf{1}_k^T - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}}\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ijk}^*\mathbf{I})\mathbf{1}_q^T\right)\mathbf{X}_i = \sum_{i\in\mathcal{L}_b}\mathbf{g}_i. \end{aligned} \tag{3.27}$$

Then, by sharing the subgradient vector $\mathbf{g}_i$ with other BSs via inter-BS communications and assuming no error is involved in sharing $\mathbf{g}_i$, each BS i can compute the global subgradient $\mathbf{g}$ locally and updates the global intercell coupling vector $\mathbf{p}$ as follows

$$\mathbf{p}^{[t+1]} = \left[\mathbf{p}^{[t]} - \frac{\alpha\mathbf{g}^{[t]^T}}{\sqrt{t}\,\|\mathbf{g}^{[t]}\|}\right]^+, \tag{3.28}$$

where $[.]^+$ indicates the projection onto nonnegative orthant, t represents the iteration index and $\alpha > 0$ is the step size.

The steps of iteratively learning $\mathbf{p}$ and solving problem in (3.4) at individual BSs are summarized in Algorithm 3.2.1 and illustrated in Fig. 3.2. At each iteration t, each BS i individually solves its own subproblem in (3.19) based on the value of $\mathbf{p}$ learned from the previous iteration, obtains the subgradient vector $\mathbf{g}_i$ in accordance with (3.27) and shares it among all other BSs via inter-BS communications. Upon learning the subgradient vector $\{\mathbf{g}_j\}_{j\neq i}$, each BS i calculates the global subgradient $\mathbf{g}$ locally and updates the global coupling vector $\mathbf{p}$ according to (3.28).

Algorithm 3.2.1 can be interpreted as a learning based ICI regularization strategy, where the cooperative BSs gradually learn the ICI imposed from other BSs and iteratively attain their own beamforming solutions until a consensus on the induced ICI powers among BSs, i.e., convergence, is reached. Furthermore, Algorithm

3.2.1 is guaranteed to converge to the optimal solution of (3.4) provided a proper selection of step size α [98]. Since the solutions to (3.19) used in the intermediate iterations of Algorithm 3.2.1 are feasible beamforming vectors that satisfy the SINR constraints, the number of iterations can be limited at the cost of sub-optimal performance in order to reduce the latency and/or the signalling overhead [50].

Algorithm 3.2.1. *Distributed Algorithm for Solving* (3.12) *at individual BSs*

1: **Initialize**: $t = 0$ *and* $\mathbf{p}(0) \in \mathbb{R}^{K(N(N-1)+1)\times 1}$;
2: **repeat** *at each* BS_i
3: ***while*** *the solutions to* (3.19) *is not converged* ***do***
4: *Each BS locally solves its subproblem* i *in* (3.19);
5: *Each BS calculates the local subgradient* $\mathbf{g}_i$ *using* (3.25);
6: *Each BS learns* $\{\mathbf{g}_j\}_{j\neq i}$ *from other BSs via inter-BS communications;*
7: *Upon obtaining subgradient vector* $\mathbf{g}_i$ *from all other BSs, each BS locally calculates the global subgradient as* $\mathbf{g} = \sum_{i\in\mathcal{L}_b} \mathbf{g}_i$;
8: *Each BS updates the global variable* $\mathbf{p}$ *according to* (3.28);
9: *Increment the iteration index* $t = t + 1$;
10: ***end while***
11: ***if*** $\mathbf{W}_{ik}^*$ *is rank-one* ***then***
12: *The optimal* $\mathbf{w}_{ik}$ *is the eigenvector of* $\mathbf{W}_{ik}^*$;
13: ***else***
14: *Apply the standard Gaussian randomization method [99] to approximate rank-one* $\mathbf{w}_{ik}$ *solutions;*
15: ***end if***
16: ***return*** $\{\mathbf{w}_{ik}\}_{i,k}$.

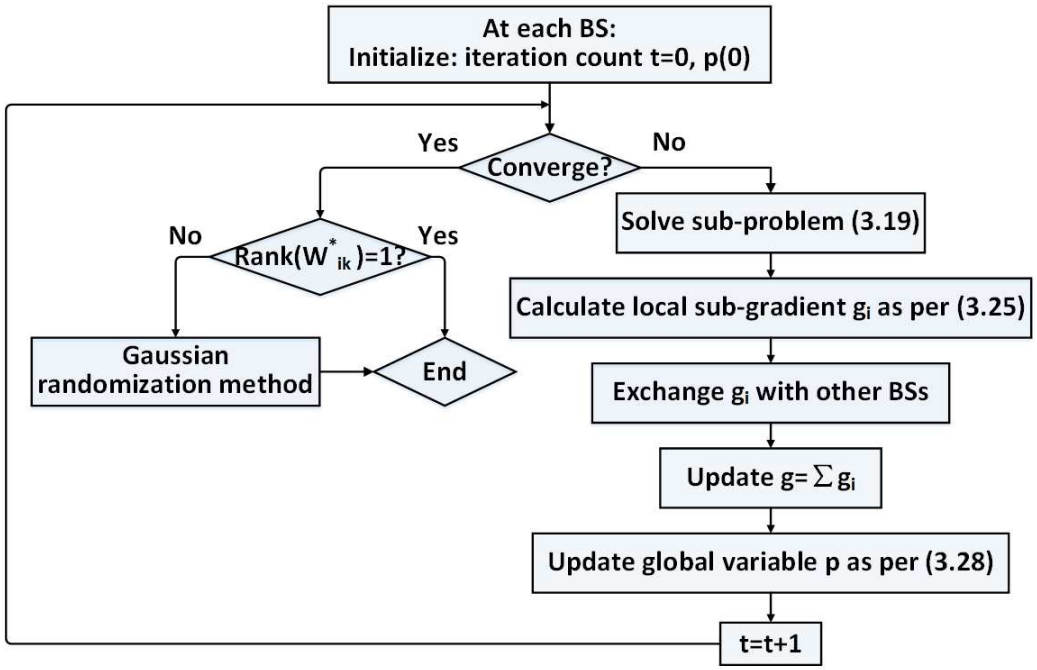


Figure 3.2: Flowchart of Algorithm 3.2.1.

3.2.4 Fronthaul Signalling Overhead and Computational Complexity Analysis

In this section, the fronthaul signaling overhead of the proposed strategy, the baseline coordinated beamforming design in [100] that requires full CSI to be shared among BSs, and the distributed robust beamforming design based on the principle of alternating direction method of multipliers (ADMM) technique in [10] will be analyzed.

For the i-th BS, the major information that need to be exchanged with the other $N-1$ BSs in each iteration of the proposed Algorithm 3.2.1 is the subgradient $\mathbf{g}_i$ that contains NK non-zero real-valued entries, i.e., $\text{tr}(\lambda_{ik}^*\mathbf{I}), \forall k$ and $\text{tr}(\lambda_{ijk}^*\mathbf{I}), \forall k, j \neq i$. The resulting inter-BS communication overhead is $O(NK(N-1))$ and thus, the total signaling overhead among all the BSs is $O(\xi N^2 K(N-1))$, where ξ is the total number of iterations of Algorithm 3.2.1. However, for the coordinated beamforming design in [100] that requires full CSI exchange, the information that need to be shared at each BS is $O(NK(N-1))$ of $M \times M$ complex-valued CSI matrices. The total

signaling overhead is then $O(4M^2N^2K(N-1))$. The ratio of fronthaul signaling load for the proposed strategy over coordinated beamforming design in [100] can be expressed as $\varphi = \frac{\xi}{4M}$. As will be evident in Section 3.4, our simulation results suggest that the proposed Algorithm 3.2.1 converges within only several iterations. Thus, with increasing number of antenna elements per BS, e.g., massive multiple-input multiple-output in 5G network, the proposed transmission strategy requires lighter inter-BS communication overhead as compared to the coordinated beamforming design that requires full CSI exchange. Interestingly, the ADMM based beamforming design in [10] requires each BS to inform other $N-1$ BSs with its NK real-valued local ICI variables at each iteration, resulting in a total fronthaul signaling load of $O(\xi N^2K(N-1))$. Hence, the ADMM design in [10] incurs a same per-iteration fronthaul signalling overhead as compared with the proposed Algorithm 3.2.1.

Next, we compare the computational complexity of the subproblem in (3.19) and the subproblem of ADMM approach in [10], in terms of number of optimization variables and number of constraints [50]. The subproblem in (3.19) has M^2K optimization variables, whereas the subproblem of the ADMM approach in [10] has $M^2K + 2NK + 1$ optimization variables. Both subproblems have NK number of LMI constraints and K number of matrix non-negativity constraints, whereas, the subproblem of the ADMM approach in [10] has additional $(N+1)K$ scalar non-negativity constraints, a quadratic constraint, and a linear constraint. Hence, Algorithm 3.2.1 has slightly lower computational complexity per subproblem as compared with the ADMM approach in [10]. Since the outputs of the intermediate iterations of the ADMM approach in [10] are not necessarily feasible for the original problem, an additional subproblem, similar to (3.19), needs to be solved at each BS to obtain feasible beamforming vectors [50]. Besides, the applicability of both Algorithm 3.2.1 and the ADMM approach in [10] are limited to rank-one solutions only, since the Gaussian randomization method [99] for approximation of higher-rank solutions do not support decentralized implementations [50].

Note that the convergence behaviour of the proposed subgradient based learning

algorithm and the ADMM approach in [10] depends on the selection of step size α and the augmented penalty parameter c, respectively. Both algorithms have similar and relatively fast convergence, provided a proper selection of α and c, respectively, e.g., $c = 50$ and $\alpha = 0.01$ [50]. Furthermore, it has been compared in [50] that the proposed algorithm with $\alpha = 0.09$ performs slightly better within the first few iterations, whereas the ADMM approach in [10] with $c = 2$ has much faster convergence at the later iterations.

3.3 Robust Transmission in Multicell Networks with Probabilistic Constraints involving Instantaneous CSI Uncertainties

Follow a similar procedure as presented in Section 3.2, this section proposes a distributed probabilistic constrained transmission strategy in multi-cell interference networks that minimizes overall downlink transmit power and provides robustness against instantaneous CSI uncertainties with different SINR outage probability levels at individual UTs.

3.3.1 System Model and Problem Formulation

Similar to Section 3.2, let us consider a multi-cell downlink network with a coordinated cluster of N cells, indexed as $\mathcal{L}_b = \{1, \cdots, N\}$. Each cell consists of a BS equipped with M antennas, transmitting to K single-antenna UTs, indexed as $\mathcal{L}_i = \{1, \cdots, K\}$, over a shared frequency band. The instantaneous channel vector from BS_i to UT_{jk}, i.e., $\mathbf{h}_{ijk} \in \mathbb{C}^{M\times 1}$, can be modelled as $\mathbf{h}_{ijk} = \mathbf{C}_{ijk}^{1/2}\mathbf{h}_w$ [101], where the entries of $\mathbf{h}_{ijk}$ are correlated, the entries of $\mathbf{h}_w$ are independent and identically distributed (i.i.d.) ZMCSCG random variables, and $\mathbf{C}_{ijk} \in \mathbb{C}^{M\times M}$ is the channel covariance matrix of UT_{jk}, as seen by the i-th BS. Without loss of generality, it is assumed that both the BSs and UTs have the prefect knowledge of $\mathbf{C}_{ijk}$, whereas

only partial information of $\mathbf{h}_w$, i.e., $\hat{\mathbf{h}}_w$, is known due to minimum mean square error (MMSE) estimation. Let the MMSE estimation error be denoted as $\mathbf{e}_w = \mathbf{h}_w - \hat{\mathbf{h}}_w$, then the true channel vector $\mathbf{h}_{ijk}$ can be modeled as

$$\mathbf{h}_{ijk} \quad = \mathbf{C}_{ijk}^{1/2}\mathbf{h}_w = \mathbf{C}_{ijk}^{1/2}(\hat{\mathbf{h}}_w + \mathbf{e}_w) = \hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk} \qquad \forall i,j,k, \tag{3.29}$$

where $\hat{\mathbf{h}}_w, \mathbf{e}_w \in \mathbb{C}^{M\times 1}$ are uncorrelated and their entries are i.i.d. ZMCSCG random variables, i.e., $[\hat{\mathbf{h}}_w]_t \sim \mathbb{CN}(0,1)$ and $[\mathbf{e}_w]_t \sim \mathbb{CN}(0,\sigma_t^2)$ [101]. $\hat{\mathbf{h}}_{ijk}$ denotes the estimated channel vector and $\mathbf{e}_{ijk}$ represents the corresponding CSI error vector. Assuming $\mathbb{E}(|s_{ik}|^2) = 1$, the SINR at UT_{ik} is then given by

$$\mathrm{SINR}_{ik} = \frac{|\mathbf{h}_{iik}^H \mathbf{w}_{ik}|^2}{\sum\limits_{\substack{n\neq k,\\ n\in\mathcal{L}_i}} |\mathbf{h}_{iik}^H \mathbf{w}_{in}|^2 + \sum\limits_{\substack{j\neq i,\ m\in\mathcal{L}_i\\ j\in\mathcal{L}_b}}\sum |\mathbf{h}_{jik}^H \mathbf{w}_{jm}|^2 + \sigma_{ik}^2}. \tag{3.30}$$

In order to optimize the overall transmit power while guaranteeing the SINR at the individual UTs with certain outage probabilities in the presence of CSI uncertainties, the following robust transmission strategy is considered, as

$$\begin{aligned} \min_{\mathbf{w}_{ik},\forall i,k} \quad & \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i} \|\mathbf{w}_{ik}\|^2 \\ \text{s.t.} \quad & \Pr\left(\mathrm{SINR}_{ik} \geq \gamma_{ik}\right) \geq 1-\rho_{ik}, \quad \forall i,k, \end{aligned} \tag{3.31}$$

where γ_{ik} is the target SINR requested by UT_{ik} and $\rho_{ik} \in (0,1)$ is the maximum SINR outage probability.

3.3.2 Optimization of Problem in (3.31)

In this section, let us start by introducing slack variables $\{p_{ijk}\}_{i,j,k} \in \mathbb{R}$ to (3.31) to account for the coupling effects among the multiple cells, as

$$\min_{\mathbf{w}_{ik}, p_{ijk}} \sum_{i \in \mathcal{L}_b} \sum_{k \in \mathcal{L}_i} \|\mathbf{w}_{ik}\|^2 \tag{3.32}$$

$$\text{s.t.} \quad \Pr\left(\frac{|\left(\hat{\mathbf{h}}_{iik} + \mathbf{e}_{iik}\right)^H \mathbf{w}_{ik}|^2}{\sum\limits_{\substack{n \neq k, \\ n \in \mathcal{L}_i}} |\left(\hat{\mathbf{h}}_{iik} + \mathbf{e}_{iik}\right)^H \mathbf{w}_{in}|^2 + \sum\limits_{\substack{l \neq i, \\ l \in \mathcal{L}_b}} p_{lik} + \sigma_{ik}^2} \geq \gamma_{ik} \right) \geq 1 - \rho_{ik}, \quad \forall i, k, \tag{3.33}$$

$$\Pr\left(\sum_{m \in \mathcal{L}_i} |\left(\hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk}\right)^H \mathbf{w}_{im}|^2 \leq p_{ijk} \right) \geq 1 - \rho_{ik}, \ \forall i, j \neq i, k, \tag{3.34}$$

where p_{ijk} indicates the ICI from BS_i to UT_{jk}. Let the rank-one positive semidefinite matrix be defined as $\mathbf{W}_{ik} = \mathbf{w}_{ik}\mathbf{w}_{ik}^H$, the set of constraints (3.34) and (3.34) can be expanded, respectively, as

$$\Pr\left(\text{tr}(-\mathbf{B}_{ik}\boldsymbol{\Delta}_{iik}) \leq \Theta + \text{tr}(\mathbf{B}_{ik}\mathbf{e}_{iik}\mathbf{e}_{iik}^H)\right) \geq 1 - \rho_{ik}, \tag{3.35}$$

$$\Pr\left(\text{tr}(\mathbf{Q}_{ijk}\boldsymbol{\Delta}_{ijk}) \leq \Upsilon - \text{tr}(\mathbf{Q}_{ijk}\mathbf{e}_{ijk}\mathbf{e}_{ijk}^H)\right) \geq 1 - \rho_{ik}, \tag{3.36}$$

where

$$\begin{cases} \mathbf{B}_{ik} = \gamma_{ik}^{-1}\mathbf{W}_{ik} - \sum\limits_{\substack{n\neq k,\\ n\in\mathcal{L}_i}} \mathbf{W}_{in}, \\ \boldsymbol{\Delta}_{iik} = \hat{\mathbf{h}}_{iik}\mathbf{e}_{iik}^H + \mathbf{e}_{iik}\hat{\mathbf{h}}_{iik}^H, \\ \Theta = \mathrm{tr}(\mathbf{B}_{ik}\hat{\mathbf{h}}_{iik}\hat{\mathbf{h}}_{iik}^H) - \sum\limits_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} - \sigma_{ik}^2, \end{cases} \tag{3.37}$$

$$\begin{cases} \mathbf{Q}_{ijk} = \sum\limits_{m\in\mathcal{L}_i} \mathbf{W}_{im}, \\ \boldsymbol{\Delta}_{ijk} = \hat{\mathbf{h}}_{ijk}\mathbf{e}_{ijk}^H + \mathbf{e}_{ijk}\hat{\mathbf{h}}_{ijk}^H, \\ \Upsilon = p_{ijk} - \mathrm{tr}(\mathbf{Q}_{ijk}\hat{\mathbf{h}}_{ijk}\hat{\mathbf{h}}_{ijk}^H). \end{cases} \tag{3.38}$$

In order to deal with the unknown terms that involve $\mathbf{e}_{iik}\mathbf{e}_{iik}^H$ and $\mathbf{e}_{ijk}\mathbf{e}_{ijk}^H$, the slack variables $\pi_1, \pi_2 \in \mathbb{R}$ are introduced and it is further assumed that the summation of error variance of each entry of $\mathbf{e}_{ijk}$ lies within a hyper-spherical region with radius of d_e, i.e., $\|\mathbf{e}_{ijk}\|^2 = \sum_{t=1}^{M} |[\mathbf{e}_{ijk}]_t|^2 \leq d_e^2$. Due to the fact that in practice, the entries of $\mathbf{e}_{ijk},\ \forall i,j,k$ are unbounded random variables, the constraints $\|\mathbf{e}_{ijk}\|^2 \leq d_e^2$ naturally indicate that the CSI errors lie within the hyper-spherical uncertainty region with a certain probability. Therefore, the radius of uncertainty region d_e should be carefully chosen in accordance with the predefined outage probability, i.e., d_e is a function of ρ_{ik}. Hence, the problem in (3.32) can be reformulated as

$$\begin{aligned} \min_{\mathbf{W}_{ik},\Theta,\Upsilon,\pi_1,\pi_2} \quad & \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\mathbf{W}_{ik}) \\ \text{s.t.} \quad & \Pr\left(\mathrm{tr}(-\mathbf{B}_{ik}\boldsymbol{\Delta}_{iik}) \leq \Theta + \pi_1\right) \geq 1-\rho_{ik}, \\ & \Pr\left(\mathrm{tr}(\mathbf{Q}_{ijk}\boldsymbol{\Delta}_{ijk}) \leq \Upsilon + \pi_2\right) \geq 1-\rho_{ik}, \\ & \mathrm{tr}(\mathbf{B}_{ik}\mathbf{e}_{iik}\mathbf{e}_{iik}^H) \geq \pi_1, \quad \forall i,k, \\ & -\mathrm{tr}(\mathbf{Q}_{ijk}\mathbf{e}_{ijk}\mathbf{e}_{ijk}^H) \geq \pi_2,\ \forall i, j\neq i, k, \\ & \|\mathbf{e}_{ijk}\|^2 \leq d_e^2(\rho_{ik}), \quad \forall i,j,k, \\ & \mathbf{W}_{ik} \succeq 0, \quad \forall i,k, \\ & \mathrm{rank}\left(\mathbf{W}_{ik}\right) = 1, \quad \forall i,k. \end{aligned} \tag{3.39}$$

The problem in (3.39) is numerically intractable since the inclusion of estimation uncertainties in SINR constraints naturally lead to an infinite number of convex sets. In the sequel, following the similar principles as in [58], the first two probabilistic constraints of the problems in (3.39) can be first equivalently converted into more convenient forms through the following Lemma.

Lemma 3.3.1. *Let $\mathbf{\Delta} \in \mathbb{C}^{M\times M}$ be a Hermitian random matrix with each ZMCSCG element being characterized as $[\mathbf{\Delta}]_{cd} \sim \mathbb{CN}(0, \sigma_{cd}^2)$. Then, for any Hermitian matrix $\mathbf{A}$, $\mathbf{A} \in \mathbb{C}^{M\times M}$,*

$$tr(\mathbf{A\Delta}) \sim \mathbb{N}(0, \|\mathcal{D}_{\mathbf{\Delta}} vec(\mathbf{A})\|^2),$$
$$tr(\mathbf{A\Delta}) = \|\mathcal{D}_{\mathbf{\Delta}} vec(\mathbf{A})\| U,\ U \sim \mathbb{N}(0,1),$$

where $\mathcal{D}_{\mathbf{\Delta}} = diag(vec(\mathbf{\Sigma_\Delta}^H))$ denotes a real-valued $M^2 \times M^2$ diagonal matrix and $\mathbf{\Sigma_\Delta}$ denotes a real-valued $M \times M$ matrix with each entry $[\mathbf{\Sigma_\Delta}]_{cd} = \sigma_{cd}$, i.e.,

$$\mathcal{D}_{\mathbf{\Delta}} = \begin{bmatrix} \sigma_{11} & 0 & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \sigma_{1M} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{21} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 & \sigma_{MM} \end{bmatrix}.$$

Proof: Please refer to a similar proof as for Lemma 3.2.1 in Appendix A.

By applying Lemma 3.3.1 and the CDF of a standard normal distribution, i.e., $\phi(u) = \Pr(U \le u) = \frac{1}{2}[1 + \text{erf}(\frac{u}{\sqrt{2}})]$, where $U \sim \mathbb{N}(0,1)$, the first and the second

probabilistic constraints in problem (3.39), respectively, can be expressed as

$$\begin{aligned} & \Pr\left(\mathrm{tr}(-\mathbf{B}_{ik}\boldsymbol{\Delta}_{iik}) \leq \Theta + \pi_1\right) \\ = \quad & \Pr\left(\|\mathcal{D}_{\boldsymbol{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik})\|U \leq \Theta + \pi_1\right) = \Pr\left(U \leq \frac{\Theta + \pi_1}{\|\mathcal{D}_{\boldsymbol{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik})\|}\right) \\ = \quad & \frac{1}{2}[1 + \mathrm{erf}\left(\frac{\Theta + \pi_1}{\sqrt{2}\|\mathcal{D}_{\boldsymbol{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik})\|}\right)] \geq 1 - \rho_{ik}, \end{aligned} \tag{3.40}$$

$$\begin{aligned} & \Pr\ \left(\mathrm{tr}(\mathbf{Q}_{ijk}\boldsymbol{\Delta}_{ijk}) \leq \Upsilon + \pi_2\right) \\ = \quad & \Pr\left(U \leq \frac{\Upsilon + \pi_2}{\|\mathcal{D}_{\boldsymbol{\Delta}_{ijk}}\mathrm{vec}(\mathbf{Q}_{ijk})\|}\right) \\ = \quad & \frac{1}{2}[1 + \mathrm{erf}\left(\frac{\Upsilon + \pi_2}{\sqrt{2}\|\mathcal{D}_{\boldsymbol{\Delta}_{ijk}}\mathrm{vec}(\mathbf{Q}_{ijk})\|}\right)] \geq 1 - \rho_{ik}, \end{aligned} \tag{3.41}$$

which are equivalent to the following expressions, respectively,

$$\sqrt{2}\mathrm{erf}^{-1}(1 - 2\rho_{ik})\|\mathcal{D}_{\boldsymbol{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik})\| \leq \Theta + \pi_1, \tag{3.42}$$

$$\sqrt{2}\mathrm{erf}^{-1}(1 - 2\rho_{ik})\|\mathcal{D}_{\boldsymbol{\Delta}_{ijk}}\mathrm{vec}(\mathbf{Q}_{ijk})\| \leq \Upsilon + \pi_2. \tag{3.43}$$

Then the first two probabilistic constraints in (3.39) can be transformed into tractable forms using Lemma 3.2.2 in Section 3.2.2. Applying Lemma 3.2.2 to (3.42) and (3.43), the first two probabilistic constraints in (3.39) can be reformulated as LMI constraints, respectively, as

$$\begin{bmatrix} \frac{\Theta+\pi_1}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\boldsymbol{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik}) \\ \mathrm{vec}^H(-\mathbf{B}_{ik})\mathcal{D}_{\boldsymbol{\Delta}_{iik}} & \frac{\Theta+\pi_1}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix} \succeq 0, \tag{3.44}$$

$$\begin{bmatrix} \frac{\Upsilon+\pi_2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\boldsymbol{\Delta}_{ijk}}\mathrm{vec}(\mathbf{Q}_{ijk}) \\ \mathrm{vec}^H(\mathbf{Q}_{ijk})\mathcal{D}_{\boldsymbol{\Delta}_{ijk}} & \frac{\Upsilon+\pi_2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix} \succeq 0. \tag{3.45}$$

However, the problem in (3.39) is still numerically intractable as terms that involve

$\mathbf{e}_{iik}\mathbf{e}_{iik}^H$ and $\mathbf{e}_{ijk}\mathbf{e}_{ijk}^H$ is unknown to the BSs. Thus, following the similar principles as in [10], the problem of intractability can be overcome via the following Lemma.

Lemma 3.3.2. *(S-procedure [8]) The implication* $\mathbf{e}^H\mathbf{A}_1\mathbf{e} + 2\Re(\mathbf{b}_1^H\mathbf{e}) + d_1 \leq 0 \Rightarrow \mathbf{e}^H\mathbf{A}_2\mathbf{e} + 2\Re(\mathbf{b}_2^H\mathbf{e}) + d_2 \leq 0$, *where* $\mathbf{A}_i \in \mathbb{H}^{M\times M}$, $\mathbf{b}_i \in \mathbb{C}^M$, $d_i \in \mathbb{R}$ *and* $\mathbf{e} \in \mathbb{C}^{M\times 1}$, *holds if and only if there exists a* $\mu \geq 0$ *such that*

$$\begin{bmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^H & d_2 \end{bmatrix} \preceq \mu \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^H & d_1 \end{bmatrix}.$$

To apply Lemma 3.3.2, let us first expand the third, fourth and fifth constraints in (3.39) in their equivalent quadratic forms of $\mathbf{e}_{iik}$ and $\mathbf{e}_{ijk}$, respectively, as

$$\begin{cases} \mathbf{e}_{iik}^H\mathbf{I}_M\mathbf{e}_{iik} - d_e^2 \leq 0, \\ -\mathbf{e}_{iik}^H\mathbf{B}_{ik}\mathbf{e}_{iik} + \pi_1 \leq 0, \ \forall i,k, \end{cases} \tag{3.46}$$

$$\begin{cases} \mathbf{e}_{ijk}^H\mathbf{I}_M\mathbf{e}_{ijk} - d_e^2 \leq 0, \\ \mathbf{e}_{ijk}^H\mathbf{Q}_{ijk}\mathbf{e}_{ijk} + \pi_2 \leq 0, \ \forall i, j \neq i,k. \end{cases} \tag{3.47}$$

Then, the constraints (3.46) and (3.47) can be rewritten in terms of LMI constraints, as

$$\begin{aligned} &\begin{bmatrix} \mathbf{B}_{ik} + \mu_{ik}\mathbf{I}_M & 0 \\ 0 & -\pi_1 - \mu_{ik}d_e^2 \end{bmatrix} \succeq 0, \\ &\mu_{ik} \geq 0, \ \forall i,k, \\ &\begin{bmatrix} -\mathbf{Q}_{ijk} + \mu_{ijk}\mathbf{I}_M & 0 \\ 0 & -\pi_2 - \mu_{ijk}d_e^2 \end{bmatrix} \succeq 0, \\ &\mu_{ijk} \geq 0, \ \forall i, j \neq i,k, \end{aligned} \tag{3.48}$$

where the set of auxiliary parameters $\mu_{ik} \geq 0$ and $\mu_{ijk} \geq 0$ appear as a result of the application of Lemma 3.3.2. Finally, combining (3.44), (3.45) with (3.48) and relaxing the set of non-convex rank-one constraints of $\text{rank}(\mathbf{W}_{ik}) = 1, \ \forall i,k$, via standard SDR approach [38], the problem in (3.39) can be reformulated as a SDP form with LMI

constraints, as

$$\begin{aligned}
&\min_{\mathbf{W}_{ik}\succeq 0,\pi_1,\pi_2,\mu_{ik},\mu_{ijk}} \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\mathbf{W}_{ik}) \qquad (3.49)\\
\text{s.t.}\quad &\begin{bmatrix} \frac{\Theta+\pi_1}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik}) \\ \mathrm{vec}^H(-\mathbf{B}_{ik})\mathcal{D}_{\mathbf{\Delta}_{iik}} & \frac{\Theta+\pi_1}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix} \succeq 0,\\
&\begin{bmatrix} \mathbf{B}_{ik}+\mu_{ik}\mathbf{I}_M & 0 \\ 0 & -\pi_1-\mu_{ik}d_e^2 \end{bmatrix} \succeq 0,\\
&\mu_{ik}\geq 0,\ \forall i,k,\\
&\begin{bmatrix} \frac{\Upsilon+\pi_2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{ijk}}\mathrm{vec}(\mathbf{Q}_{ijk}) \\ \mathrm{vec}^H(\mathbf{Q}_{ijk})\mathcal{D}_{\mathbf{\Delta}_{ijk}} & \frac{\Upsilon+\pi_2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix} \succeq 0,\\
&\begin{bmatrix} -\mathbf{Q}_{ijk}+\mu_{ijk}\mathbf{I}_M & 0 \\ 0 & -\pi_2-\mu_{ijk}d_e^2 \end{bmatrix} \succeq 0,\\
&\mu_{ijk}\geq 0,\ \forall i,j\neq i,k,
\end{aligned}$$

The problem in (3.49) can now be optimally solved in a centralized fashion. In case that the rank of optimal solutions to (3.49) are greater than one, a similar randomization method to [11] can be adopted to approximate the feasible rank-one solution. In the next section, the problem in (3.49) will be decomposed via primal decomposition [98].

3.3.3 Distributed Optimization of problem in (3.49)

Let the global intercell coupling variables $\mathbf{p}\in\mathbb{R}^{N(N-1)K\times 1}$ be defined as

$$\mathbf{p}=\Big[p_{121},p_{122},...,p_{12K},...,p_{N11},...,p_{NN-1K}\Big]^T. \qquad (3.50)$$

Then the direction vector $\mathbf{d}_{iik}$ and $\mathbf{d}_{ijk} \in \{0,1\}^{N(N-1)K\times 1}$ will be employed to extract $\sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik}$ and p_{ijk} from $\mathbf{p}$, respectively, as

$$\begin{cases} \sum\limits_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} &= \mathbf{d}_{iik}^T\mathbf{p}, \quad \forall k, \\ p_{ijk} &= \mathbf{d}_{ijk}^T\mathbf{p}, \quad \forall j\neq i,k. \end{cases} \tag{3.51}$$

Similar to Section 3.2.3, the problem in (3.49) can be decomposed into two levels of optimization, where for any given $\mathbf{p}$, N sub-problems can be individually solved at each BS i, as

$$\begin{aligned} \min_{\mathbf{W}_{ik},\pi_1,\pi_2,\mu_{ik},\mu_{ijk}} \quad & f_i(\mathbf{W}_{ik}) \triangleq \sum_{k\in\mathcal{L}_i} \mathrm{tr}(\mathbf{W}_{ik}) \\ \text{s.t.} \quad & \mathbf{T}_{ik} = \mathbf{T}'_{ik} - (\mathbf{d}_{iik}^T\mathbf{p})\mathbf{I}_{(M^2+1)} \succeq 0, \\ & \mathbf{E}_{ik} = \begin{bmatrix} \mathbf{B}_{ik}+\mu_{ik}\mathbf{I}_M & 0 \\ 0 & -\pi_1-\mu_{ik}d_e^2 \end{bmatrix} \succeq 0, \\ & \mu_{ik}\geq 0,\ \forall i,k, \\ & \mathbf{T}_{ijk} = \mathbf{T}'_{ijk} + (\mathbf{d}_{ijk}^T\mathbf{p})\mathbf{I}_{(M^2+1)} \succeq 0, \\ & \mathbf{E}_{ijk} = \begin{bmatrix} \mu_{ijk}\mathbf{I}_M-\mathbf{Q}_{ijk} & 0 \\ 0 & -\pi_2-\mu_{ijk}d_e^2 \end{bmatrix} \succeq 0, \\ & \mu_{ijk}\geq 0,\ \forall i,j\neq i,k, \\ & \mathbf{W}_{ik}\succeq 0, \end{aligned} \tag{3.52}$$

where

$$\mathbf{T}'_{ik} = \begin{bmatrix} \frac{\mathrm{tr}(\mathbf{B}_{ik}\hat{\mathbf{h}}_{iik}\hat{\mathbf{h}}_{iik}^H)-\sigma_{ik}^2+\pi_1}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{iik}}\mathrm{vec}(-\mathbf{B}_{ik}) \\ \mathrm{vec}^H(-\mathbf{B}_{ik})\mathcal{D}_{\mathbf{\Delta}_{iik}} & \frac{\mathrm{tr}(\mathbf{B}_{ik}\hat{\mathbf{h}}_{iik}\hat{\mathbf{h}}_{iik}^H)-\sigma_{ik}^2+\pi_1}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix}, \tag{3.53}$$

$$\mathbf{T}'_{ijk} = \begin{bmatrix} \frac{-\mathrm{tr}(\mathbf{Q}_{ijk}\hat{\mathbf{h}}_{ijk}\hat{\mathbf{h}}_{ijk}^H)+\pi_2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})}\mathbf{I}_{M^2} & \mathcal{D}_{\mathbf{\Delta}_{ijk}}\mathrm{vec}(\mathbf{Q}_{ijk}) \\ \mathrm{vec}^H(\mathbf{Q}_{ijk})\mathcal{D}_{\mathbf{\Delta}_{ijk}} & \frac{-\mathrm{tr}(\mathbf{Q}_{ijk}\hat{\mathbf{h}}_{ijk}\hat{\mathbf{h}}_{ijk}^H)+\pi_2}{\sqrt{2}\mathrm{erf}^{-1}(1-2\rho_{ik})} \end{bmatrix},$$

and the master problem for updating the global variable $\mathbf{p}$, is defined as $\min_{\mathbf{p}} \sum_{i\in\mathcal{L}_b} f_i^*(\mathbf{p})$. Follow a similar procedure as presented in Section 3.2.3, let $\lambda_{ik}, \lambda_{ijk} \in \mathbb{H}^{(M^2+1)\times(M^2+1)}$, $\alpha_{ik}, \alpha_{ijk} \in \mathbb{H}^{(M+1)\times(M+1)}$ and $\beta_{ik}, \beta_{ijk} \in \mathbb{R}$ be defined as the Lagrange multipliers, then the Lagrangian of the i-th subproblem in (3.52) for a fixed value of $\mathbf{p}$, can be expressed as

$$L_i(\{\mathbf{W}_{ik}, \lambda_{ik}, \alpha_{ik}, \beta_{ik}\}_k, \{\lambda_{ijk}, \alpha_{ijk}, \beta_{ijk}\}_{j\neq i,k}) = \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\mathbf{W}_{ik}\right) - \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\lambda_{ik}\mathbf{T}_{ik}\right)$$
$$- \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\lambda_{ijk}\mathbf{T}_{ijk}\right) - \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\alpha_{ik}\mathbf{E}_{ik}\right) - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\alpha_{ijk}\mathbf{E}_{ijk}\right) - \beta_{ik}\mu_{ik} - \beta_{ijk}\mu_{ijk}. \tag{3.54}$$

Since the problem in (3.52) is convex and satisfies the Slaters condition, strong duality holds [8] and the dual function is given by

$$\ell_i(\mathbf{p}) = \inf_{\mathbf{W}_{ik}\succeq 0} L_i = \Xi\left(\left\{\lambda_{ik}^*, \alpha_{ik}^*, \beta_{ik}^*\right\}_k, \left\{\lambda_{ijk}^*, \alpha_{ijk}^*, \beta_{ijk}^*\right\}_{j\neq i,k}\right) \tag{3.55}$$
$$+ \left(\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ik}\mathbf{I})\mathbf{d}_{iik}^T - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ijk}\mathbf{I})\mathbf{d}_{ijk}^T\right)\mathbf{p},$$

where

$$\Xi\left(\left\{\lambda_{ik}^*, \alpha_{ik}^*, \beta_{ik}^*\right\}_k, \left\{\lambda_{ijk}^*, \alpha_{ijk}^*, \beta_{ijk}^*\right\}_{j\neq i,k}\right) = \inf_{\mathbf{W}_{ik}\succeq 0} \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\mathbf{W}_{ik}\right) - \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\alpha_{ik}\mathbf{E}_{ik}\right)$$
$$- \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\alpha_{ijk}\mathbf{E}_{ijk}\right) - \beta_{ik}\mu_{ik} - \beta_{ijk}\mu_{ijk} - \sum_{k\in\mathcal{L}_i} \mathrm{tr}\big(\lambda_{ik}\mathbf{T}'_{ik}\big) - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}\big(\lambda_{ijk}\mathbf{T}'_{ijk}\big). \tag{3.56}$$

Then we can write

$$f_i^*(\mathbf{W}_{ik}^*, \mathbf{p}) = f_i^*(\mathbf{p}) = \ell_i^*(\mathbf{p}) = \mathbf{g}_i\mathbf{p} + \Xi\left(\left\{\lambda_{ik}^*, \alpha_{ik}^*, \beta_{ik}^*\right\}_k, \left\{\lambda_{ijk}^*, \alpha_{ijk}^*, \beta_{ijk}^*\right\}_{j\neq i,k}\right), \tag{3.57}$$

where

$$\mathbf{g}_i = \sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ik}^*\mathbf{I})\mathbf{d}_{iik}^T - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ijk}^*\mathbf{I})\mathbf{d}_{ijk}^T, \tag{3.58}$$

$\mathbf{g}_i \in \mathbb{R}^{1\times N(N-1)K}$ is the subgradient vector of $\ell_i^*(\mathbf{p})$ and $f_i^*(\mathbf{p})$ obtained for the i-th subproblem [98]. The global subgradient $\sum_{i\in\mathcal{L}_b} f_i^*(\mathbf{p})$, obtained for the general problem in (3.49) at a given $\mathbf{p}$, can be calculated as

$$\begin{aligned}\mathbf{g} \quad &= \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ik}^*\mathbf{I})\mathbf{d}_{iik}^T - \sum_{i\in\mathcal{L}_b}\sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}}\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ijk}^*\mathbf{I})\mathbf{d}_{ijk}^T \\ &= \sum_{i\in\mathcal{L}_b}\left(\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ik}^*\mathbf{I})\mathbf{d}_{iik}^T - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}}\sum_{k\in\mathcal{L}_i} \mathrm{tr}(\lambda_{ijk}^*\mathbf{I})\mathbf{d}_{ijk}^T\right) = \sum_{i\in\mathcal{L}_b}\mathbf{g}_i.\end{aligned} \tag{3.59}$$

Then, Algorithm 3.2.1 in Section 3.2.3 can be adopted by individual BSs to solve the problem in (3.31) distributively, where the cooperative BSs gradually learn to achieve a reasonable consensus on the global ICI.

3.4 Simulation Results

Let us consider 3 adjacent cells, each cell consists of a BS with inter-BS distance of 500 m. As shown in Fig. 3.3, 2 UTs are randomly scheduled in the vicinity of the boundaries in each cell to account for the worst ICI effect. Similar to [1], the (m,n)-th element of the channel covariance matrix $\mathbf{C}_{ijk} \in \mathbb{C}^{M\times M}$ is modeled as

$$[\mathbf{C}_{ijk}]_{mn} = e^{j\frac{2\pi\delta}{\lambda}\left[(n-m)sin\theta_{ijk}\right]} e^{-2\left[\frac{\pi\delta\sigma_a}{\lambda}(n-m)cos\theta_{ijk}\right]^2},\ m,n\in[1,M]\,, \tag{3.60}$$

where $\delta = \lambda/2$ is the spacing between two adjacent antenna elements, λ is the carrier wavelength, $\sigma_a = 2^\circ$ is angular offset standard deviation and θ_{ijk} is the angle of departure for UT_{jk} with respect to the broadside of the antenna of BS_i. Besides,

Table 3.1: Simulation parameters [1, 2]

Parameter	Value
Number of cells (N)	3
Number of UTs per cell (K)	2
Number of antennas per BS (M)	8
Distance between two adjacent BSs	500 m
Array antenna gain	15 dBi
Noise power spectral density (all users)	-174 dBm/Hz
Noise figure at user receiver	5 dB
Path loss model over a distance of ℓ m	$34.53 + 38 \log_{10}(\ell)$
Angular offset standard deviation σ_a	2°
Log-normal shadowing standard deviation σ_s	10 dB

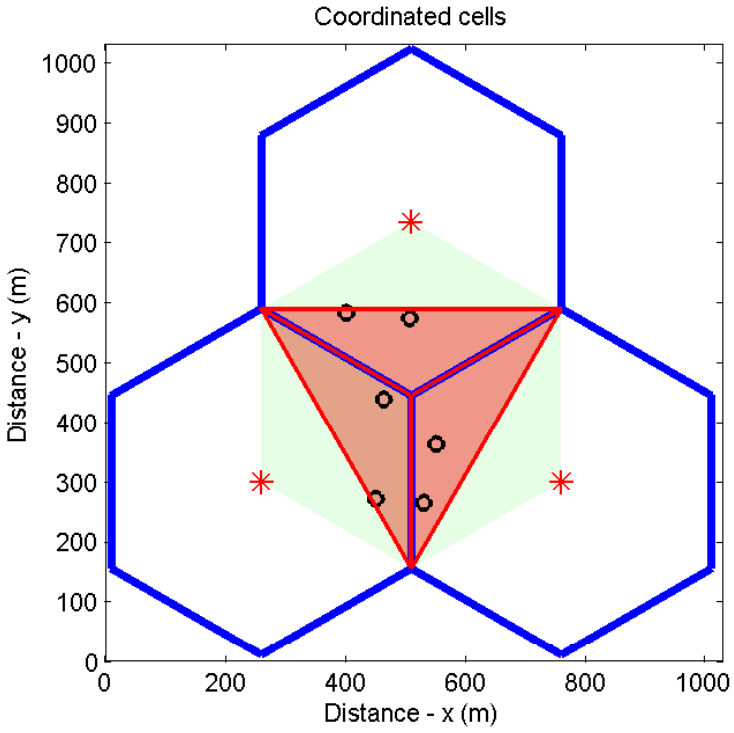


Figure 3.3: An example of user distribution in a 3-cell network.

to take consideration of path loss, shadowing and fading, the channel covariance matrix $\mathbf{C}_{ijk}$ and its corresponding random error matrix $\mathbf{\Delta}_{ijk}$ in Section 3.2, as well as the channel vector $\hat{\mathbf{h}}_{ijk}$ and its corresponding estimation error $\mathbf{e}_{ijk}$ in Section 3.3, are scaled by $G_a L_{ijk} \sigma_F^2 e^{-0.5\frac{(\sigma_s ln10)^2}{100}}$ [1], where $G_a = 15$ dBi is array antenna gain, $L_{ijk} = 34.53{+}38\log_{10}(\ell)$ represents the path loss model over a distance of ℓ m between BS_i and UT_{jk} [2], σ_F^2 is the variance of the complex Gaussian fading coefficient, $\sigma_s = 10$ dB is log-normal shadowing standard deviation and flat-fading channels are assumed. Other important parameters are presented in Table 3.1 [1, 2]. The step size in Algorithm 3.2.1 is selected as $\alpha = \frac{1}{\sqrt{t}}$ [50]. Equal SINR targets γ and equal SINR outage probability ρ are assumed for all UTs in different cells. The performance of the proposed transmission strategy is evaluated and averaged via the existing solvers, e.g., CVX [35]. The results are presented in comparison with the distributed worst-case sum-power minimization designs in [10] and [50] that provide robustness against bounded CSI error, and an outage probability based robust beamforming design based on Bernstein-type inequality method against instantaneous CSI error in [9].

It is further assumed that each entry of error matrix $\mathbf{\Delta}_{ijk}$ in Section 3.2 has the same variance $\sigma_{cd}^2 = \sigma_e^2$, whilst each entry of estimation error $\mathbf{e}_w$ in Section 3.3 has the same variance $\sigma_t^2 = \sigma^2$, i.e., $[\mathbf{e}_w]_t \sim \mathbb{CN}(0, \sigma^2)$. In the sequel, a connection between the radius of uncertainty region d_e and the outage probability ρ will be illustrated. Since $\mathbf{e}_{ijk} \in \mathbb{C}^{M\times 1}$ consists of M ZMCSCG random variables, which is equivalent to $2M$ real normal random variables, i.e., $[\mathbf{e}_{ijk}]_t = \Re\{[\mathbf{e}_{ijk}]_t\} + \Im\{[\mathbf{e}_{ijk}]_t\}$, where $\Re\{[\mathbf{e}_{ijk}]_t\} = \frac{\sigma_t}{\sqrt{2}}U,\ \Im\{[\mathbf{e}_{ijk}]_t\} = \frac{\sigma_t}{\sqrt{2}}U,\ U \sim \mathbb{N}(0,1)$, then, it can be written as

$$\begin{aligned} \|\mathbf{e}_{ijk}\|^2 &= \sum_{t=1}^{M} |[\mathbf{e}_{ijk}]_t|^2 = \sum_{t=1}^{M} (\Re([\mathbf{e}_{ijk}]_t)^2 + \Im([\mathbf{e}_{ijk}]_t)^2) \\ &= \sum_{t=1}^{2M} \frac{\sigma_t^2}{2} U^2 = \frac{\sigma^2}{2} \sum_{t=1}^{2M} U^2 \leq d_e^2(\rho). \end{aligned} \tag{3.61}$$

Then according to the definition of the CDF of chi-square distribution [102], the CDF

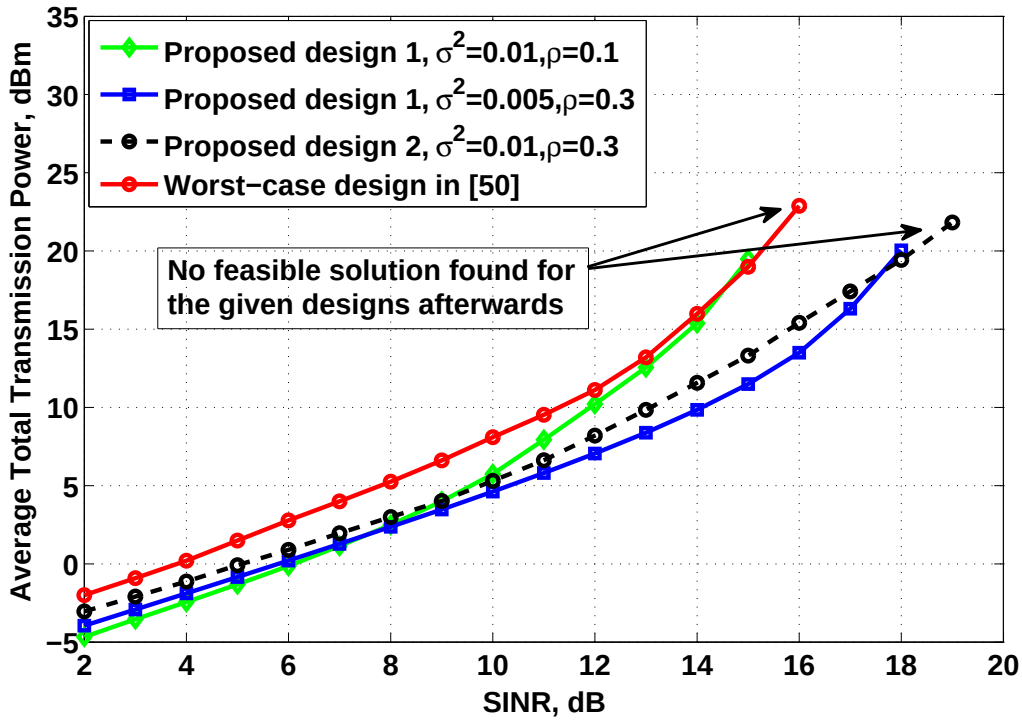
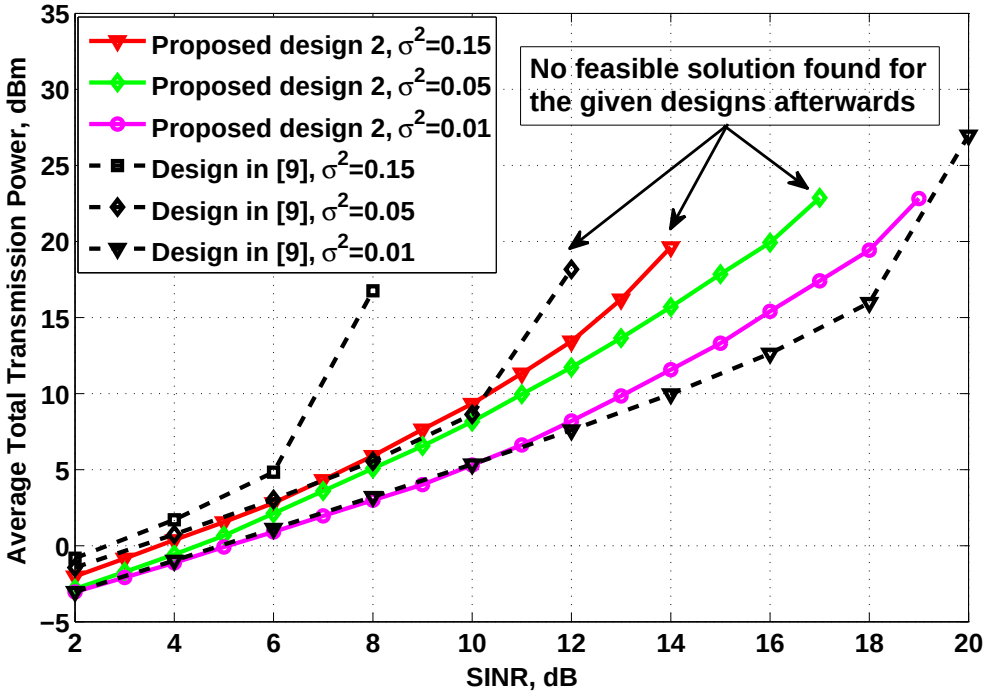


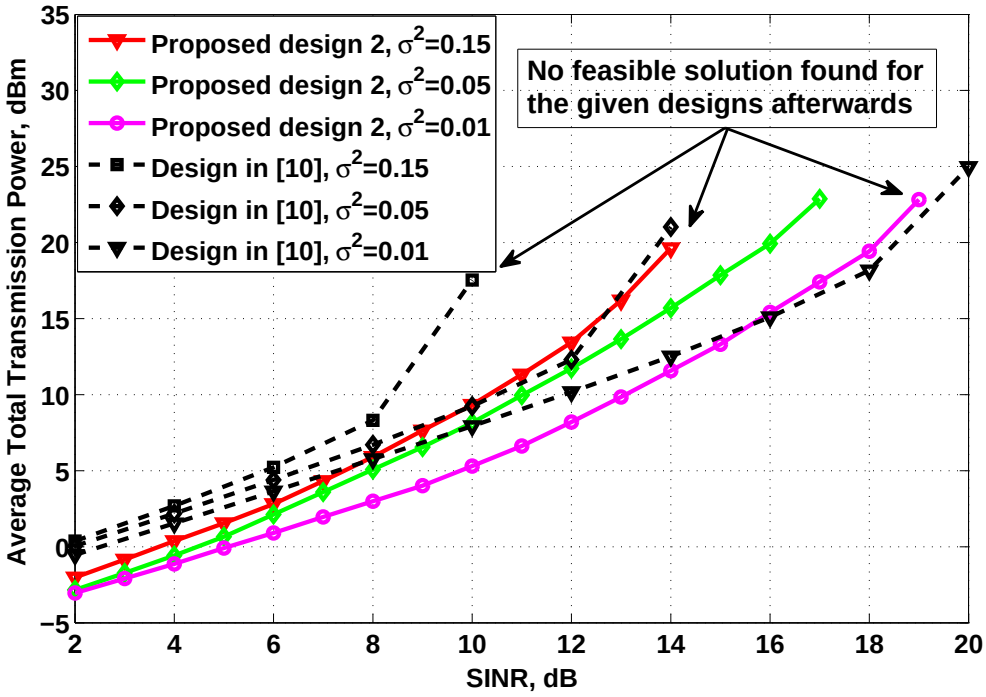
Figure 3.4: Comparison of total transmit power versus various SINR outage probabilities and error variances.

of $\Pr(\sum_{t=1}^{2M} U^2) \leq \frac{2d_e^2}{\sigma^2}$ can be expressed as $\psi_{\chi^2_{2M}}(\frac{2d_e^2}{\sigma^2}) = 1-\rho$, which indicates the probability of $1-\rho$ that a hyper-spherically bounded uncertainty region holds for radius $d_e = \sqrt{\frac{\sigma^2 \psi^{-1}_{\chi^2_{2M}}(1-\rho)}{2}}$, where $\psi^{-1}_{\chi^2_{2M}}(.)$ is the inverse CDF of a standard chi-square distribution with $2M$ degrees of freedom.

The performance comparison in terms of total transmit power of the outage probability based strategies proposed in Section 3.2 and in Section 3.3 with different SINR outage levels, against worst-case bounded error design in [50] that corresponds to $\rho = 0.1$ and $\sigma^2 = 0.01$, is presented in Fig. 3.4. It can be observed from the figure that in terms of providing better power efficiency, the strategy proposed in Section 3.2 has a performance improvement of approximately 5% as compared to the worst-case design in [50] up to medium SINR operational range. Furthermore, the strategy proposed in Section 3.2 is more power efficient than the strategy in Section 3.3 up to medium SINR operational range, whereas for higher SINR targets, the strategy in



(a)



(b)

Figure 3.5: Comparison of total transmit power with $\rho = 0.3$ for the proposed strategy and a) outage probability based design in [9], b) ADMM approach in [10].

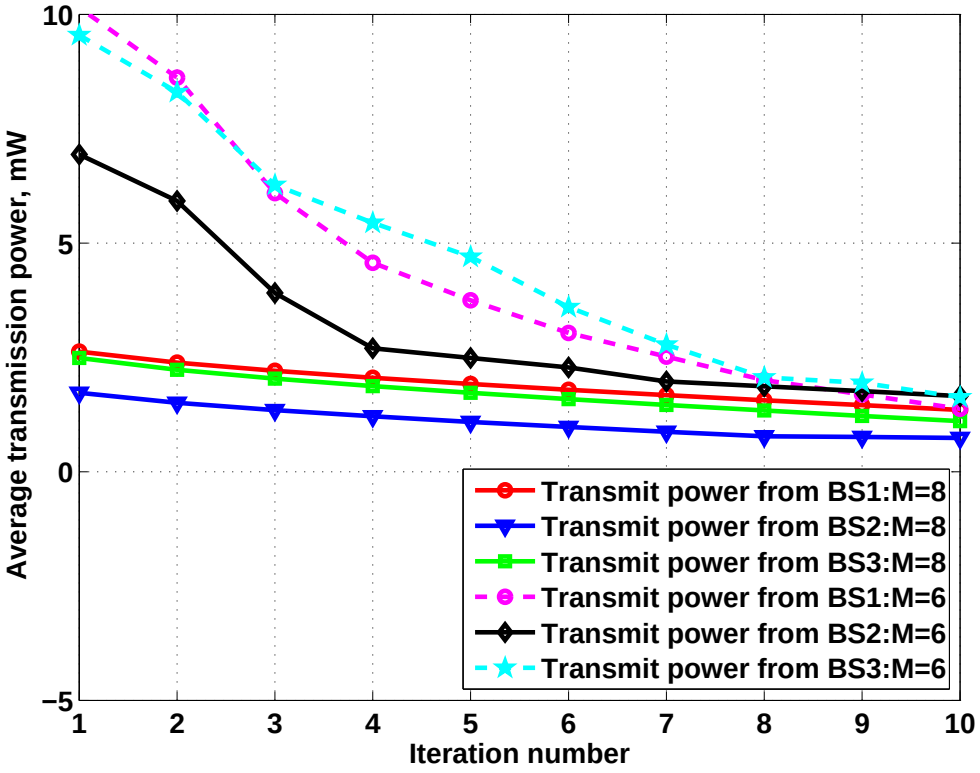


Figure 3.6: Power variation of Algorithm 3.2.1 at $\gamma = 10$ dB target SINR for $M = 6, 8$ antenna elements per BS.

Section 3.3 requests less total transmit power. One can also conclude that for a given CSI uncertainty variance, the total transmit power consumption increases with the decreasing outage probability ρ. The performance gap can be interpreted that the higher level of robustness against CSI uncertainties comes at the cost of increment in total transmit power. On the contrary, with a fixed SINR outage probability, the transmission strategy with smaller value of CSI error variance consumes less total transmit power.

Fig. 3.5 presents the performance comparison of total transmit power for the strategy proposed in Section 3.3 at $\rho = 0.3$ with different CSI error variances against an outage probability based design in [9] and the bounded error robust ADMM approach in [10]. One can conclude from the figure that the proposed strategy outperforms the designs in [10] and [9] in terms of expanding SINR operational range for the observed error variance except for the case of $\sigma^2 = 0.01$. This confirms the improved resilience against higher variance level of CSI uncertainties of the proposed strategy. In the case of $\sigma^2 = 0.01$, the proposed strategy requires approximately 5% less transmit power as compared with the conservative worst-case design in [10] for low and medium SINR operational range and closely follows the outage probability based design in [9] up to medium target SINR.

The power variation of proposed Algorithm 3.2.1 with $\sigma_e^2 = 0.005$ and $\rho = 0.3$ at $\gamma = 10$ dB target SINR is presented in Fig. 3.6 for $M = 6, 8$ number of antenna elements per BS. It can be observed from the figure that with the increasing number of antenna elements per BS, the required transmit power at initial iteration increases significantly while the convergence speed decreases. Furthermore, the range of power variations between the initial and the final iterations decreases as we increase the number of per-BS antenna elements since extra degree and more accurate coordination can be provided by the BSs.

3.5 Concluding Remarks

In this chapter, two outage probability based distributed robust coordinated transmission strategies for minimizing the overall transmit power in downlink multi-cell interference networks in the presence of imperfect CSI are proposed. The problems are constrained to SINR requirements and provide robustness against, respectively, the statistical and instantaneous CSI uncertainties with different SINR outage probability levels at individual UTs. The numerically intractable problems are first converted into their centralized SDP forms with LMI constraints based on CDF of standard normal distribution, Schur complement, S-procedure and SDR technique. Then the general problems are decomposed into a set of parallel subproblems to be solved at individual BSs via subgradient learning iterations to coordinate the cross-link interference across the BSs with a light fronthaul signaling overhead. Simulation results confirm the advantages of the proposed strategies in terms of providing larger SINR operational range as compared with worst-case robust beamforming designs in [10, 50] and outage probability based robust beamforming design in [9]. Furthermore, in terms of power efficiency, the proposed strategies have approximately 5% performance improvement as compared to the worst-case designs in [50] and [10] up to medium SINR operational range.

Chapter 4

An UCB Algorithm for Worst-Case Distributed Robust Transmission in Multicell Networks

4.1 Introduction

This chapter introduces a robust approach for maximizing the weighted signal-to-interference-plus-noise-ratio (SINR) requirements at user terminals (UTs) in the presence of imperfect channel state information (CSI) in decentralized multicell interference networks. The optimization problem is constrained to strict available power budget at individual base stations (BSs). Based on the inverse relationship between the max-min SINR problem and the sum-power minimization problem, the original numerically intractable problem is first reformulated in an equivalent overall transmit power minimization problem constrained by a set of robust SINR constraints in the centralized worst-case scenario for a fixed SINR weight. Then, the multicell-wise centralized sum-power minimization problem for a given SINR weight is transformed into a numerically tractable form via S-procedure and semidefinite relaxation (SDR) techniques, and then decomposed into a set of independent subproblems at individual BSs. Finally, an upper confidence bound (UCB) based algorithm is introduced to distributively update SINR weights and scale the SINR targets based on individual BS power budgets, and coordinate intercell interference (ICI) among BSs with a light inter-BS communication overhead.

4.1.1 Main Contributions

The main contributions of this chapter are summarized as follows.

- In contrast to the simple bisection algorithm for updating SINR weights where

BSs individually search for their own parameter without considering other BSs, this chapter proposes an UCB algorithm for individual BSs to optimally scale their SINR targets across the involved multi-cells in a distributed manner with a light inter-BS communications overhead based on individual BS power budgets.

- The original problem formulation naturally leads to computationally intractability which is dealt with in this chapter by reformulating the original problem in its alternative tractable form. However, the reformulation adds non-convex rank-one constraints to the alternative optimization problem. Thus, firstly, the rank-one constraints are relaxed via SDR technique to find tractable solutions, and then, the solutions to the reformulated tractable problem are analytically proved to be always rank-one. Therefore, no computationally expensive randomization technique is required to find the rank-one solutions.

Simulation results confirm the advantage of the proposed strategy in terms of providing larger SINR operation range against robust distributed beamforming design in [50], as it optimally scales the SINR targets based on per BS power budgets and always provides a feasible solution at the scaled SINR target.

4.1.2 Organization

The rest of this chapter is organized as follows. Section 4.2 introduces the system model and problem formulation, where the original problem is converted to an equivalent dual problem. In Section 4.3, the intractable centralized power minimization problem is first transformed into a numerically tractable one. Then, a learning based UCB algorithm is proposed for decoupling the problem into distributed subproblems, followed by the signalling overhead and computational complexity analysis in Section 4.3.3. Simulation results are presented and analyzed in Section 4.4. Finally, Section 4.5 summarizes the chapter.

4.2 System Model and Problem Formulation

Let us consider a multi-cell downlink network with a cluster of N cells over a shared bandwidth. Each cell consists of one BS equipped with M antennas, cooperating at beamforming level and transmitting to its own K single-antenna UTs. Let BS_i, $i \in \mathcal{L}_b = \{1, \cdots, N\}$ and UT_{ik}, $k \in \mathcal{L}_i = \{1, \cdots, K\}$ represent the i-th BS and the k-th UT in cell i, respectively. Also let s_{ik} denote the data symbol for UT_{ik} and n_{ik} be the additive white Gaussian noise with variance σ_{ik}^2, $\mathbf{w}_{ik} \in \mathbb{C}^{M\times 1}$ be the associated beamforming vector and $\mathbf{h}_{ijk} \in \mathbb{C}^{M\times 1}$ represent the channel vector from BS_i to UT_{jk}. Then the signal received by UT_{ik} is given by

$$z_{ik} = \mathbf{h}_{iik}^H \mathbf{w}_{ik} s_{ik} + \sum_{\substack{n\neq k,\\ n\in\mathcal{L}_i}} \mathbf{h}_{iik}^H \mathbf{w}_{in} s_{in} + \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{m\in\mathcal{L}_i} \mathbf{h}_{jik}^H \mathbf{w}_{jm} s_{jm} + n_{ik}. \tag{4.1}$$

Let $\hat{\mathbf{h}}_{ijk} \in \mathbb{C}^{M\times 1}$ and $\mathbf{e}_{ijk} \in \mathbb{C}^{M\times 1}$, respectively, denote the estimated channel vector and the corresponding CSI perturbation vector. Then, the true channel vector $\mathbf{h}_{ijk}$ can be modeled as

$$\mathbf{h}_{ijk} = \hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk}, \quad \forall i, j, k, \tag{4.2}$$

where CSI errors are assumed to be bounded within an elliptic uncertainty region, i.e., $\mathbf{e}_{ijk}^H \mathbf{R}_{ijk} \mathbf{e}_{ijk} \leq 1$, $\forall i, j, k$, and $\mathbf{R}_{ijk} \succ 0$ specifies the shape and size of the ellipsoid. Without loss of generality, let us assume $\mathbb{E}(|s_i|^2) = 1$. Then, SINR at UT_{ik} can be formulated as

$$\text{SINR}_{ik} = \frac{|\mathbf{h}_{iik}^H \mathbf{w}_{ik}|^2}{\sum\limits_{\substack{n\neq k,\\ n\in\mathcal{L}_i}} |\mathbf{h}_{iik}^H \mathbf{w}_{in}|^2 + \sum\limits_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum\limits_{m\in\mathcal{L}_i} |\mathbf{h}_{jik}^H \mathbf{w}_{jm}|^2 + \sigma_{ik}^2}. \tag{4.3}$$

Let us consider the robust problem of maximizing the minimum weighted SINR targets at UTs in a multi-cell network subject to a set of strict upper limits on the transmit power constraints at individual BSs, e.g., due to regulation, in the presence

of CSI errors, as

$$\max_{\mathbf{w}_{ik}, c_i,} \min_{\forall i,k} \quad c_i \tag{4.4a}$$

$$\text{s.t.} \quad \text{SINR}_{ik} \geq c_i \gamma_{ik}, \quad \forall i,k,$$

$$\sum_{k \in \mathcal{L}_i} \|\mathbf{w}_{ik}\|^2 \leq P_i, \quad \forall i, \tag{4.4b}$$

$$\mathbf{e}_{ijk}^H \mathbf{R}_{ijk} \mathbf{e}_{ijk} \leq 1, \quad \forall i,j,k, \tag{4.4c}$$

where γ_{ik} is the SINR requirement at UT_{ik} and P_i represents the available power budget at the i-th BS that can not be relaxed. The introduction of an auxiliary variable c_i is to lower bound the worst-case scaled SINR, which indicates the percentage coefficient of the desired SINR targets that can be satisfied at UTs as a result of strict power constraints at BS i. In fact, the aim of the proposed optimization is to maximize the worst-case achievable SINR targets at UTs subject to strict limitations on transmit power at individual BSs. Contrary to the sum power minimization approach, e.g., [50], problem (4.4) always admits a feasible solution at scaled SINR and is more flexible since it can be used to determine whether, in a power-constrained system, a specified set of SINR targets can be satisfied or not [40].

Since problem (4.4) is numerically intractable due to the coupling effects among BSs operating under unit frequency bandwidth as well as the robust constraints against CSI uncertainties, let us begin by introducing an alternative overall transmit power minimization problem at BS i, as

$$\min_{\mathbf{w}_{ik}, \forall k} \quad f_i(\mathbf{w}_{ik}) \triangleq \sum_{k \in \mathcal{L}_i} \|\mathbf{w}_{ik}\|^2 \tag{4.5}$$

$$\text{s.t.} \quad \text{SINR}_{ik} \geq c_i \gamma_{ik}, \ \forall i,k,$$

$$\mathbf{e}_{ijk}^H \mathbf{R}_{ijk} \mathbf{e}_{ijk} \leq 1, \ \forall k.$$

Note that following similar procedures as in Section 3.2.3, for any fixed value of c_i, the alternative power minimization problem in (4.5) can be solved in a similar way as for the subproblem in (3.19) within any individual cell i distributively.

In the sequel, the optimal solutions to problems in (4.4) and (4.5) within any cell i for a given SINR weight c_i, will be related through Lemma 4.2.1. Let $\mathbf{\Gamma}_i = \{\gamma_{ik}\}_k$ be a set of K target SINRs for UTs in cell i. For a given set of channels and noise powers, problem (4.4) is parameterized by $\mathbf{\Gamma}_i$ and P_i, whereas problem (4.5) is parameterized by $\mathbf{\Gamma}_i$. The dependence is captured by notations $s(\mathbf{\Gamma}_i, P_i)$ and $f(\mathbf{\Gamma}_i)$, respectively. Also, let $c_i = s^*(\mathbf{\Gamma}_i, P_i)$ and $P_i = f^*(\mathbf{\Gamma}_i)$ represent, respectively, the optimal values, i.e., maximum worst-case scaled SINR and the minimum power, of problems (4.4) and (4.5).

Lemma 4.2.1. *Problem* (4.4) *and problem* (4.5) *are inverse problems and are related as follows:*

$$c_i = s^*(\mathbf{\Gamma}_i, f^*(c_i\mathbf{\Gamma}_i)),$$
$$P_i = f^*(s^*(\mathbf{\Gamma}_i, P_i)\mathbf{\Gamma}_i).$$

Proof: See [40] and [103].

Thus, considering c_i as a variable of optimization, the optimal solutions to (4.4) can be obtained in an approximate manner via alternating between solving problem (4.5) for a fixed c_i, and searching over different c_i based on per BS power restriction.

4.3 Distributed Optimization of Problem (4.4)

It has been proved in [40] that for a single-cell multicasting network, the optimality of solution for max-min SINR problem can be guaranteed by alternatively solving power minimization problem for a fixed c_i and applying a simple bisection search over c_i. In a multi-cell scenario, however, the obtained c_i may not be globally optimum if BSs individually search for their own c_i without considering other BSs. Consequently, following similar steps as in Section 3.2.3, the distributed optimization of problem (4.5) for a fixed c_i will be first introduced in Section 4.3.1, and an UCB algorithm will be introduced in Section 4.3.2 to search for the optimal c_i across all BSs in a decentralized fashion.

4.3.1 Distributed Optimization of (4.5) for a Fixed c_i

Let us start by introducing a centralized formulation of the total transmit power optimization problem in (4.5) for a fixed value of c_i to account for the coupling effects among the BSs. Introducing slack variables $\{p_{ijk}\}_{i,j,k} \in \mathbb{R}$ to indicate ICI from BS_i to UT_{jk}, problem (4.5) can be generalized as

$$\min_{\mathbf{w}_{ik}, p_{ijk}} \quad \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i} \|\mathbf{w}_{ik}\|^2$$

$$\text{s.t.} \quad \frac{|\left(\hat{\mathbf{h}}_{iik}+\mathbf{e}_{iik}\right)^H \mathbf{w}_{ik}|^2}{\sum\limits_{\substack{n\neq k,\\ n\in\mathcal{L}_i}} |\left(\hat{\mathbf{h}}_{iik}+\mathbf{e}_{iik}\right)^H \mathbf{w}_{in}|^2 + \sum\limits_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} + \sigma_{ik}^2} \geq c_i\gamma_{ik}, \quad \forall i,k, \tag{4.6a}$$

$$p_{ijk} \geq \sum_{m\in\mathcal{L}_i} |\left(\hat{\mathbf{h}}_{ijk}+\mathbf{e}_{ijk}\right)^H \mathbf{w}_{im}|^2, \quad \forall i, j\neq i, k, \tag{4.6b}$$

$$\mathbf{e}_{ijk}^H \mathbf{R}_{ijk} \mathbf{e}_{ijk} \leq 1, \quad \forall i,j,k. \tag{4.6c}$$

Let the rank-one positive semidefinite matrix be defined as $\mathbf{W}_{ik} = \mathbf{w}_{ik}\mathbf{w}_{ik}^H$, the constraints in (4.6a) and (4.6b) can be rewritten as

$$\left(\hat{\mathbf{h}}_{iik}+\mathbf{e}_{iik}\right)^H \mathbf{\Phi}_{ik} \left(\hat{\mathbf{h}}_{iik}+\mathbf{e}_{iik}\right) \geq \sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} + \sigma_{ik}^2, \forall i,k \tag{4.7}$$

$$p_{ijk} \geq \left(\hat{\mathbf{h}}_{ijk}+\mathbf{e}_{ijk}\right)^H \mathbf{\Psi}_{ijk} \left(\hat{\mathbf{h}}_{ijk}+\mathbf{e}_{ijk}\right), \forall i, j\neq i, k, \tag{4.8}$$

where $\mathbf{\Phi}_{ik} = (c_i\gamma_{ik})^{-1}\mathbf{W}_{ik} - \sum_{\substack{n\neq k,\\ n\in\mathcal{L}_i}} \mathbf{W}_{in}$ and $\mathbf{\Psi}_{ijk} = \sum_{m\in\mathcal{L}_i} \mathbf{W}_{im}$. Hence, problem (4.6) can be reformulated as

$$\begin{aligned}
\min_{\mathbf{W}_{ik}, p_{ijk}} \quad & \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\mathbf{W}_{ik}\right) \qquad (4.9)\\
\text{s.t.} \quad & \left(\hat{\mathbf{h}}_{iik} + \mathbf{e}_{iik}\right)^H \mathbf{\Phi}_{ik} \left(\hat{\mathbf{h}}_{iik} + \mathbf{e}_{iik}\right) \geq \sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} + \sigma_{ik}^2, \quad \forall i,k\\
& p_{ijk} \geq \left(\hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk}\right)^H \mathbf{\Psi}_{ijk} \left(\hat{\mathbf{h}}_{ijk} + \mathbf{e}_{ijk}\right), \quad \forall i, j\neq i, k,\\
& \mathbf{e}_{ijk}^H \mathbf{R}_{ijk} \mathbf{e}_{ijk} \leq 1, \quad \forall i,j,k\\
& \mathbf{W}_{ik} \succeq 0, \quad \forall i,k,\\
& \mathrm{rank}\left(\mathbf{W}_{ik}\right) = 1, \quad \forall i,k.
\end{aligned}$$

The set of non-convex rank-one constraints in problem (4.9) can be relaxed via SDR approach [38]. However, it is still numerically intractable as the remaining robust SINR constraints that involve bounded CSI errors have to be satisfied in the intersection of infinite number of convex sets. Following the similar principles as in [10], the intractability can be overcome via Lemma 3.3.2, i.e., S-Procedure, in Section 3.3.2.

Let the constraints in (4.9) be expanded in their equivalent quadratic forms of $\mathbf{e}_{iik}$ and $\mathbf{e}_{ijk}$, respectively, as

$$\begin{aligned}
& \mathbf{e}_{iik}^H \mathbf{R}_{iik} \mathbf{e}_{iik} - 1 \leq 0 \;\Rightarrow \qquad (4.10)\\
& -\mathbf{e}_{iik}^H \mathbf{\Phi}_{ik} \mathbf{e}_{iik} - (\mathbf{\Phi}_{ik}\hat{\mathbf{h}}_{iik})^H \mathbf{e}_{iik} - \mathbf{e}_{iik}^H \mathbf{\Phi}_{ik} \hat{\mathbf{h}}_{iik} - v_{ik} \leq 0, \; \forall i,k,\\
& \mathbf{e}_{ijk}^H \mathbf{R}_{ijk} \mathbf{e}_{ijk} - 1 \leq 0 \;\Rightarrow \qquad (4.11)\\
& \mathbf{e}_{ijk}^H \mathbf{\Psi}_{ijk} \mathbf{e}_{ijk} + (\mathbf{\Psi}_{ijk}\hat{\mathbf{h}}_{ijk})^H \mathbf{e}_{ijk} + \mathbf{e}_{ijk}^H \mathbf{\Psi}_{ijk} \hat{\mathbf{h}}_{ijk} - v'_{ijk} \leq 0, \; \forall i, j\neq i, k,
\end{aligned}$$

where $v_{ik} = \hat{\mathbf{h}}_{iik}^H \mathbf{\Phi}_{ik} \hat{\mathbf{h}}_{iik} - \sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} - \sigma_{ik}^2$ and $v'_{ijk} = -\hat{\mathbf{h}}_{ijk}^H \mathbf{\Psi}_{ijk} \hat{\mathbf{h}}_{ijk} + p_{ijk}$. Applying Lemma 3.3.2 to (4.10) and (4.11), problem in (4.9) can be rewritten in semidefinite

programming form with linear matrix inequality constraints, as

$$\begin{aligned}
\min_{\mathbf{W}_{ik},v,{v'}_{ijk},\mu_{ik},\mu_{ijk}} \quad & \sum_{i\in\mathcal{L}_b}\sum_{k\in\mathcal{L}_i}\mathrm{tr}\left(\mathbf{W}_{ik}\right) \\
\text{s.t.} \quad & \begin{bmatrix} \mu_{ik}\mathbf{R}_{iik}+\mathbf{\Phi}_{ik} & \mathbf{\Phi}_{ik}\hat{\mathbf{h}}_{iik} \\ (\mathbf{\Phi}_{ik}\hat{\mathbf{h}}_{iik})^H & -\mu_{ik}+v_{ik} \end{bmatrix} \succeq 0, \\
& \mu_{ik}\geq 0,\ \forall i,k, \\
& \begin{bmatrix} \mu_{ijk}\mathbf{R}_{ijk}-\mathbf{\Psi}_{ijk} & -\mathbf{\Psi}_{ijk}\hat{\mathbf{h}}_{ijk} \\ (-\mathbf{\Psi}_{ijk}\hat{\mathbf{h}}_{ijk})^H & -\mu_{ijk}+{v'}_{ijk} \end{bmatrix} \succeq 0, \\
& \mu_{ijk}\geq 0,\ \forall i,j\neq i,k, \\
& \mathbf{W}_{ik}\succeq 0,\ \forall i,k,
\end{aligned} \tag{4.12}$$

where the set of auxiliary parameters $\mu_{ik}\geq 0$ and $\mu_{ijk}\geq 0$ appear as a result of the application of Lemma 3.3.2. The convex optimization problem in (4.12) can now be solved in a centralized fashion. In the sequel, problem in (4.12) will be decomposed via primal decomposition [98]. Let us define $\mathbf{p}\in\mathbb{R}^{(N(N-1)K)\times 1}$ as a real-valued vector that contains the global intercell coupling variables, i.e., $\mathbf{p} = \left[p_{121},p_{122},...,p_{12K},...,p_{N11},...,p_{NN-1K}\right]^T$. Then, $\sum_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik}$ and p_{ijk} can be extracted from global intercell coupling variable $\mathbf{p}$ by using direction vectors $\mathbf{d}_{iik}$ and $\mathbf{d}_{ijk}\in\{0,1\}^{(N(N-1)K)\times 1}$, respectively, as

$$\begin{cases} \sum\limits_{\substack{l\neq i,\\ l\in\mathcal{L}_b}} p_{lik} &= \mathbf{d}_{iik}^T\mathbf{p}, \quad \forall k, \\ p_{ijk} &= \mathbf{d}_{ijk}^T\mathbf{p}, \quad \forall j\neq i,k. \end{cases} \tag{4.13}$$

According to decomposition theory [98] and following similar procedure as in Section 3.2.3, the problem in (4.12) can be decomposed into N sub-problems $f_i(\mathbf{W}_{ik})$ at individual BSs for a fixed global variable $\mathbf{p}$, and a master problem $\min\limits_{\mathbf{p}}\sum\limits_{i\in\mathcal{L}_b} f_i^*(\mathbf{p})$ for updating the global variable $\mathbf{p}$. Consequently, for any given $\mathbf{p}$, the sub-problem

at any BS i can be expressed as

$$\begin{aligned}
\min_{\mathbf{W}_{ik},\mu_{ik},\mu_{ijk}} \quad & f_i(\mathbf{W}_{ik}) \triangleq \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\mathbf{W}_{ik}\right) \\
\text{s.t.} \quad & \mathbf{E}_{ik} = \mathbf{E}'_{ik} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{d}_{iik}^T\mathbf{p} \end{bmatrix} \succeq 0, \\
& \mathbf{F}_{ijk} = \mathbf{F}'_{ijk} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{d}_{ijk}^T\mathbf{p} \end{bmatrix} \succeq 0, \\
& \mu_{ik} \geq 0, \ \forall k, j \neq i, \\
& \mu_{ijk} \geq 0, \ \forall k, j \neq i, \\
& \mathbf{W}_{ik} \succeq 0, \ \forall k,
\end{aligned} \tag{4.14}$$

where

$$\begin{aligned}
\mathbf{E}'_{ik} &= \begin{bmatrix} \mu_{ik}\mathbf{R}_{iik} + \mathbf{\Phi}_{ik} & \mathbf{\Phi}_{ik}\hat{\mathbf{h}}_{iik} \\ (\mathbf{\Phi}_{ik}\hat{\mathbf{h}}_{iik})^H & \hat{\mathbf{h}}_{iik}^H\mathbf{\Phi}_{ik}\hat{\mathbf{h}}_{iik} - \sigma_{ik}^2 - \mu_{ik} \end{bmatrix}, \\
\mathbf{F}'_{ijk} &= \begin{bmatrix} \mu_{ijk}\mathbf{R}_{ijk} - \mathbf{\Psi}_{ijk} & -\mathbf{\Psi}_{ijk}\hat{\mathbf{h}}_{ijk} \\ (-\mathbf{\Psi}_{ijk}\hat{\mathbf{h}}_{ijk})^H & -\hat{\mathbf{h}}_{ijk}^H\mathbf{\Psi}_{ijk}\hat{\mathbf{h}}_{ijk} - \mu_{ijk} \end{bmatrix}.
\end{aligned} \tag{4.15}$$

Lemma 4.3.1. *The optimal solutions to the problems* (4.14) *satisfy rank* $(\mathbf{W}_{ik}^*) = 1$ *with probability one.*

Proof: Please refer to the Appendix B. Let $\lambda_{ik}, \lambda_{ijk} \in \mathbb{H}^{(M+1)\times(M+1)}$, $\mathbf{A}_{ik} \in \mathbb{H}^{M\times M}$ and $\beta_{ik}, \beta_{ijk} \in \mathbb{R}$ be defined as Lagrange multipliers, then the Lagrangian of the i-th subproblem in (4.14) can be expressed as

$$\begin{aligned}
&L_i(\{\mathbf{A}_{ik}, \lambda_{ik}, \beta_{ik}\}_k, \{\lambda_{ijk}, \beta_{ijk}\}_{j\neq i,k}) = \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\mathbf{W}_{ik}\right) - \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\lambda_{ik}\mathbf{E}_{ik}\right) \\
&- \sum_{\substack{j\neq i, \\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \mathrm{tr}\left(\lambda_{ijk}\mathbf{F}_{ijk}\right) - \beta_{ik}\mu_{ik} - \beta_{ijk}\mu_{ijk} - \mathbf{A}_{ik}\mathbf{W}_{ik}.
\end{aligned} \tag{4.16}$$

Since the problem in (4.14) is convex and satisfies the Slater condition, strong

duality holds [8] and the dual function is given by

$$\ell_i(\mathbf{p}) = \inf_{\mathbf{W}_{ik}\succeq 0} L_i = \Xi\left(\{\lambda_{ik},\beta_{ik}\}_k, \{\lambda_{ijk},\beta_{ijk}\}_{j\neq i,k}\right) + \left(\sum_{k\in\mathcal{L}_i}[\lambda_{ik}]_{(M+1)(M+1)}\,\mathbf{d}_{iik}^T - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}}\sum_{k\in\mathcal{L}_i}[\lambda_{ijk}]_{(M+1)(M+1)}\,\mathbf{d}_{ijk}^T\right)\mathbf{p}, \tag{4.17}$$

where

$$\Xi_i\left(\{\lambda_{ik},\beta_{ik}\}_k, \{\lambda_{ijk},\beta_{ijk}\}_{j\neq i,k}\right) = \inf_{\mathbf{W}_{ik}\succeq 0}\sum_{k\in\mathcal{L}_i}\mathrm{tr}\left(\mathbf{W}_{ik}\right) - \sum_{k\in\mathcal{L}_i}\mathrm{tr}\big(\lambda_{ik}\mathbf{E}'_{ik}\big) - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}}\sum_{k\in\mathcal{L}_i}\mathrm{tr}\big(\lambda_{ijk}\mathbf{F}'_{ijk}\big) - \mathbf{A}_{ik}\mathbf{W}_{ik}. \tag{4.18}$$

Defining $\mathbf{g}_i \in \mathbb{R}^{1\times(N(N-1)K)}$ as

$$\mathbf{g}_i = \sum_{k\in\mathcal{L}_i}[\lambda^*_{ik}]_{(M+1)(M+1)}\,\mathbf{d}_{iik}^T - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}}\sum_{k\in\mathcal{L}_i}\left[\lambda^*_{ijk}\right]_{(M+1)(M+1)}\mathbf{d}_{ijk}^T, \tag{4.19}$$

then we can write

$$f_i^*(\mathbf{W}_{ik}^*) = f_i^*(\mathbf{p}) = \ell_i^*(\mathbf{p}) = \mathbf{g}_i\mathbf{p} + \Xi_i\left(\{\lambda^*_{ik},\beta^*_{ik}\}_k, \left\{\lambda^*_{ijk},\beta^*_{ijk}\right\}_{j\neq i,k}\right). \tag{4.20}$$

It can be easily concluded from (4.20) that for any given $\hat{\mathbf{p}}$, the following inequality holds

$$\begin{aligned}\ell_i^*(\mathbf{p}) &= \mathbf{g}_i\mathbf{p} + \Xi_i\left(\{\lambda^*_{ik},\beta^*_{ik}\}_k, \left\{\lambda^*_{ijk},\beta^*_{ijk}\right\}_{j\neq i,k}\right) \\ &= \mathbf{g}_i(\mathbf{p}-\hat{\mathbf{p}}) + \mathbf{g}_i\hat{\mathbf{p}} + \Xi_i\left(\{\lambda^*_{ik},\beta^*_{ik}\}_k, \left\{\lambda^*_{ijk},\beta^*_{ijk}\right\}_{j\neq i,k}\right) \leq \mathbf{g}_i(\mathbf{p}-\hat{\mathbf{p}}) + \ell_i^*(\hat{\mathbf{p}}).\end{aligned} \tag{4.21}$$

Hence, $\mathbf{g}_i$ is the subgradient vector of $\ell_i^*(\mathbf{p})$ and $f_i^*(\mathbf{p})$. Following a similar sequence of analysis as for the sub-problem in (4.14), one can easily verify that the subgradient

of the general problem in (4.12), i.e., $\sum_{i\in\mathcal{L}_b} f_i^*(\mathbf{p})$, at a given value of $\mathbf{p}$, denoted by $\mathbf{g} \in \mathbb{R}^{1\times(N(N-1)K)}$, can be calculated as

$$\mathbf{g} = \sum_{i\in\mathcal{L}_b}\Big(\sum_{k\in\mathcal{L}_i} \left[\lambda_{ik}^*\right]_{(M+1)(M+1)} \mathbf{d}_{iik}^T - \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \left[\lambda_{ijk}^*\right]_{(M+1)(M+1)} \mathbf{d}_{ijk}^T\Big) = \sum_{i\in\mathcal{L}_b} \mathbf{g}_i. \quad (4.22)$$

To achieve minimization of total transmit power across multiple cells for a fixed c_i while optimally account for the coupling intercell effects in a distributed manner, we proceed as follows. At a given value of c_i, each BS i individually solves its subproblem (4.14), obtains its subgradient vector $\mathbf{g}_i$ and shares it with other BSs via an inter-BS communications phase. Then, each BS i locally calculates the global subgradient $\mathbf{g}$ as per (4.22) and updates the global coupling vector $\mathbf{p}$ via projected subgradient learning iterations, as follows,

$$\mathbf{p}^{[t+1]} = \max\left(0, \mathbf{p}^{[t]} - \frac{\alpha \mathbf{g}^{[t]^T}}{\sqrt{t}\,\|\mathbf{g}^{[t]}\|}\right), \quad (4.23)$$

where the superscript ^{t} denotes the iteration index of inner problem (4.14) and α represents the step size. The steps are summarized in Algorithm 4.3.1.

As mentioned in the beginning of Section 4.3, simply applying a one-dimensional bisection search over c_i for distributed approach may not yield a global optimal solution for c_i since each BS will find its own c_i individually without considering other BSs. Consequently, let us consider searching for the global optimal c_i as a multi-armed bandit (MAB) problem and propose a reinforcement learning based UCB algorithm in the sequel to search for the optimal c_i across all BSs in a decentralized fashion.

4.3.2 UCB Algorithm for Finding the Globally Optimal c_i

The MAB problem is formulated as a system of N arms, each being associated with i.i.d. stochastic rewards. The objective is to maximize the accumulated reward by alternatively acquiring new knowledge, known as exploration, while simultaneously optimizing the decisions based on existing partial knowledge, known as exploitation,

in multiple rounds [12].

This chapter extracts an abstract idea of MAB problem, where playing an arm at each round is equivalent to running Algorithm 4.3.1, i.e., *Exploration for finding reward of the i-th BS*, to estimate the reward for a BS at the n-th round. In the sequel, an UCB Algorithm, i.e., Algorithm 4.3.2, will be introduced to search for the global optimal c_i at the i-th BS, as shown in Fig. 4.1. Due to the fact that the coupling effect among all BSs is negligible for low SINR targets, each BS individually searching for their own c_i barely induces interference to other BSs. Thus, Algorithm 4.3.2 first executes coarse tuning to adjust c_i rapidly so that the actual transmit power at each BS is close to the per-BS power limitation, as per Step 2 of Algorithm 4.3.2. Then, by adopting fine tuning, BSs alternatively adjust their c_i on the basis of their rewards and interactions. Let $\bar{\mathcal{R}}(\mathrm{BS}_i^{[n]})$ and $\hat{\mathcal{R}}(\mathrm{BS}_i^{[n]})$, respectively, be defined as the estimated mean reward and adjusted reward for the i-th BS at the n-th round. In the n-th round of fine tuning, each BS calculates the estimated mean reward as per Algorithm 4.3.1 and the adjusted reward as per Step 5 of Algorithm 4.3.2. Then, in the $(n+1)$-th round, only the BSs with the highest adjusted reward will run the Algorithm 4.3.1 to search for a new c_i, while other BSs will maintain the same c_i as in the previous round. Note that $\sqrt{\frac{3\ln(n)}{2T_i^{[n]}}}$ in Algorithm 4.3.2 reflects the fundamental trade-off between exploration that examines the unknown rewards and exploitation that chooses the best-possible rewards so far, where $T_i^{[n]}$ denotes the total number of times the Algorithm 4.3.1 has been run at the i-th BS in the n-th round.

By adjusting the value of ξ, $c_i^{[\min]}$ and $c_i^{[\max]}$, one can control the overall system performance conveniently. Furthermore, the UCB algorithm can be used to determine the exact level of under- or over-satisfaction of SINR targets, provided that a proper searching interval of c_i is selected, i.e., $c_i^{[\min]}$ and $c_i^{[\max]}$ [40]. For instance, by setting $c_i^{[\min]} = 0$ and $c_i^{[\max]} = 1$, Algorithm 4.3.2 is equivalent to an sum power minimization approach, but can always provide a feasible solution at scaled SINR. Whereas if no limit is set to $c_i^{[\max]}$, Algorithm 4.3.2 will provide optimal solutions to problem (4.4) with inequality power constraint (4.4b) being met with equality.

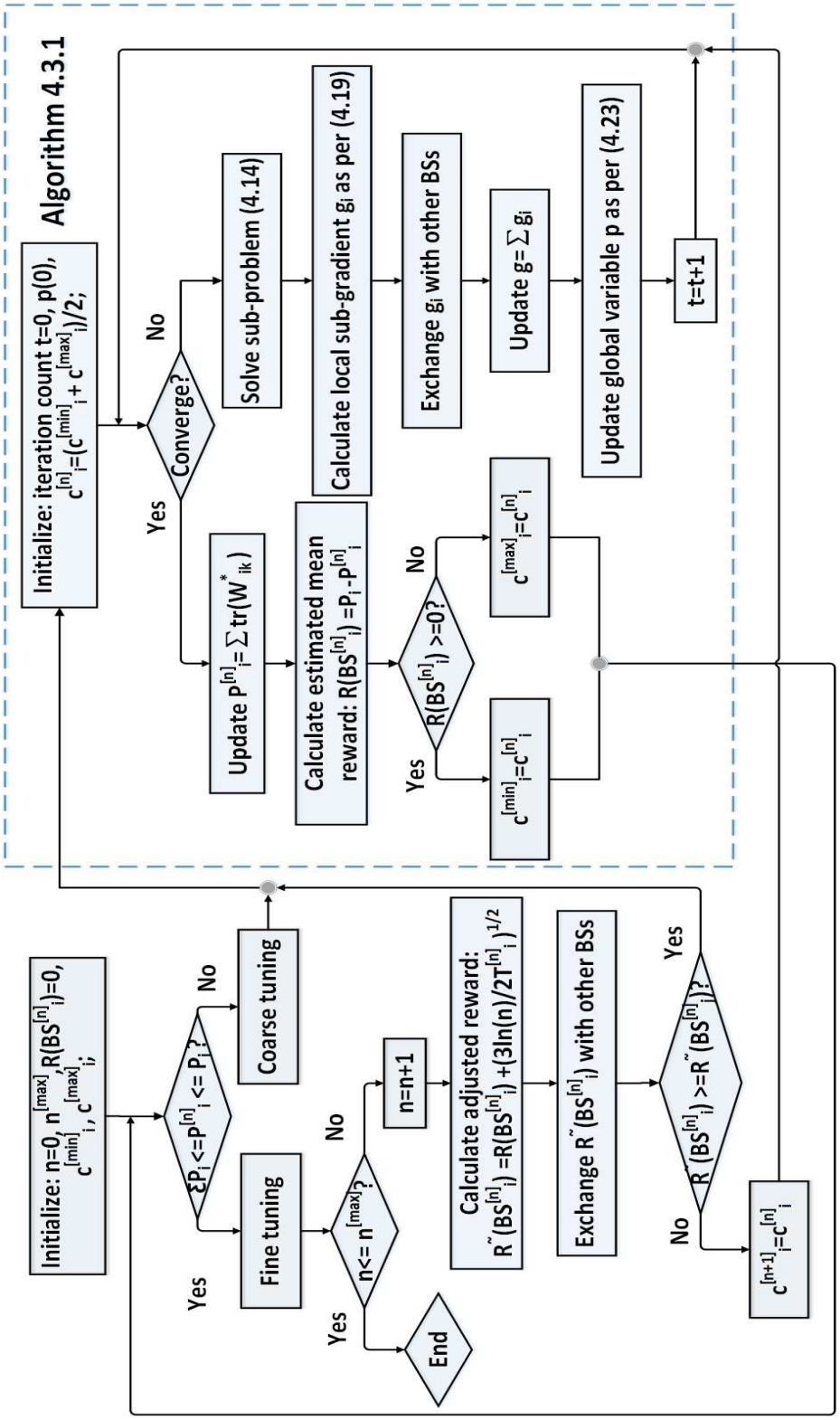


Figure 4.1: Flowchart diagram of the proposed UCB algorithm

Algorithm 4.3.1. *Exploration for finding reward of the i-th BS*

1: **Initialize**: $t = 0$, $\mathbf{p}(0) \in \mathbb{R}^{K(N(N-1)+1)\times 1}$;
2: $c_i^{[n]} = (c_i^{[min]} + c_i^{[max]})/2$;
3: ***while*** *the inner problem in* (4.14) *is not converged* ***do***
4: *Solve* (4.14);
5: *Calculate the local subgradient* $\mathbf{g}_i$ *using* (4.19);
6: *Exchange* $\mathbf{g}_i$ *with the other BSs;*
7: *Form the global subgradient as* $\mathbf{g} = \sum_{i\in\mathcal{L}_b} \mathbf{g}_i$;
8: *Update the global variable* $\mathbf{p}$ *according to* (4.23);
9: *Increment the iteration number* $t = t + 1$;
10: ***end while***
11: $\hat{P}_i^{[n]} = f_i^*(c_i^{[n]}\mathbf{\Gamma}_i) = \sum_{k\in\mathcal{L}_i} tr(\mathbf{W}_{ik}^*)$;
12: *Calculate estimated mean reward* $\bar{\mathcal{R}}(BS_i^{[n]}) = P_i - \hat{P}_i^{[n]}$;
13: **if** $\bar{\mathcal{R}}(BS_i^{[n]}) \geq 0$;
14: **then** $c_i^{[min]} = c_i^{[n]}$;
15: **else** $c_i^{[max]} = c_i^{[n]}$;
16: **end if**

Algorithm 4.3.2. *UCB Algorithm for finding global optimal* c_i

1: **Initialize**: $n = 0$, $\bar{\mathcal{R}}(BS_i^{[n]}) = \hat{\mathcal{R}}(BS_i^{[n]}) = 0$, n_{max}, $c_i^{[min]}$, $c_i^{[max]}$;
2: **Coarse tuning**: *Run Algorithm 4.3.1 until* $\hat{P}_i^{[n]} \in [\xi P_i \ P_i],\ 0 \leq \xi \leq 1$;
3: **Fine tuning**: **While** $n \leq n_{max}$ **do**
4: $n = n + 1$;
5: *Calculate the adjusted reward* $\hat{\mathcal{R}}(BS_i^{[n]}) = \bar{\mathcal{R}}(BS_i^{[n]}) + \sqrt{\frac{3ln(n)}{2T_i^{[n]}}}$;
6: BS_i *exchanges* $\hat{\mathcal{R}}(BS_i^{[n]})$ *with other BSs;*
7: **if** $\hat{\mathcal{R}}(BS_i^{[n]}) \geq \hat{\mathcal{R}}(BS_j^{[n]}),\ \forall j \in \mathcal{L}_b, j \neq i$
8: **then** *Run Algorithm 4.3.1;*
9: **else** $c_i^{[n+1]} = c_i^{[n]}$ *and run line 3-11 of Algorithm 4.3.1;*
10: **end while**
11: ***return*** $\{\mathbf{w}_{ik}\}_{i,k}$ *and* c_i

4.3.3 Fronthaul Signaling Overhead and Computational Complexity Analysis

In this section, the per iteration fronthaul signaling overhead as well as the per subproblem computational complexity of the proposed strategy will be analyzed and compared against the alternating direction method of multipliers (ADMM) approach in [10]. The proposed strategy requires NK non-zero real-valued entries, i.e., $[\lambda^*_{ik}]_{(M+1)(M+1)}, \forall k$ and $[\lambda^*_{ijk}]_{(M+1)(M+1)}, \forall k, j \neq i$, for the i-th BS to exchange with other BSs in each iteration t. The resulting inter-BS communication overhead per iteration for all BSs is $O(N^2K(N-1))$, and the total signalling overhead of the proposed strategy is $O(\omega\xi N^2K(N-1))$, where ξ is the total number of iterations of Algorithm 4.3.1 and ω is the total iteration number of Algorithm 4.3.2. Whereas in ADMM approach in [10], NK real-valued local ICI variables need to be informed by each BS at each iteration, resulting in a same per iteration fronthaul signaling load of $O(N^2K(N-1))$ as the proposed strategy.

In the sequel, the computational complexity of the subproblem in (4.14) and the subproblem of ADMM approach in [10] will be compared in terms of number of optimization variables and constraints. The subproblem in (4.14) has $M^2K + NK + 1$ optimization variables, whereas the subproblem in [10] has $M^2K+2NK+1$ optimization variables. Both subproblems have NK number of LMI constraints, K number of matrix non-negativity constraints, K scalar non-negativity constraints and a linear constraint. The subproblem of ADMM approach in [10], nevertheless, has additional NK scalar non-negativity constraints and a quadratic constraint. Therefore, Algorithm 4.3.1 has slightly lower computational complexity per subproblem as compared to the ADMM approach in [10].

4.4 Simulation Results

Let us consider a cluster of $N = 3$ neighbouring cells with BSs cooperating at beamforming level. $K = 2$ UTs are randomly dropped in the vicinity of the

boundaries in each cell to account for the worst coupling effect amongst BSs. Such 3-cell network is also adopted in [10] and [9], as the cell-edge UTs can benefit most from a coordinated cluster of 3 BSs. Similar to [1], a correlated channel model is adopted as $\hat{\mathbf{h}}_{ijk} = \mathbf{C}_{ijk}^{1/2}\mathbf{h}_w, \ \forall i,j,k$, where $\mathbf{h}_w \sim \mathbb{CN}(0,1) \in \mathbb{C}^{M\times 1}$. The (m,n)-th element of channel covariance matrix $\mathbf{C}_{ijk} \in \mathbb{C}^{M\times M}$ is given by $[\mathbf{C}_{ijk}]_{mn} = \sqrt{G_a L_{ijk}\sigma_F^2 e^{-0.5\frac{(\sigma_s ln10)^2}{100}}} e^{j\frac{2\pi\delta}{\lambda}\left[(n-m)sin\theta_{ijk}\right]} e^{-2\left[\frac{\pi\delta\sigma_a}{\lambda}(n-m)cos\theta_{ijk}\right]^2}, m,n \in [1,M]$ [1], where $L_{ijk} = 128.1 + 37.6\log_{10}(\ell), \ell$ in km, is the path loss between BS_i and UT_{jk} [2], σ_F^2 denotes the variance of the complex Gaussian fading coefficient, δ is the antenna spacing, λ denotes the wavelength of the carrier and θ_{ijk} is the estimated angle of departure. Equal noise variance $\sigma_{ik}^2 = -127$ dBm and SINR targets γ are used for all UTs and same per-BS transmit power restriction $P_i = 30$ dBm is applied to all BSs. The simulation parameters are summarized in Table 4.1 [1–3]. It is further assumed that the CSI errors are spherically bounded, i.e., $\mathbf{R}_{ijk} = 1/r_e^2\mathbf{I}$, with uncertainty radius of $r_e = 0.05$ for simplicity [10]. Simulation results are obtained and averaged via CVX [35]. In order to compare the proposed strategy with other energy-efficient beamforming designs, let us set $c_i^{[\max]} = 1$ and $c_i^{[\min]} = 0$ in Algorithm 4.3.2 to optimize the trade-off between power constraints at individual BSs and desired SINR targets at UTs. The comparative designs are, respectively, the conventional non-coordinated beamforming design, the centralized non-robust

Table 4.1: Simulation parameters [1–3]

Parameter	**Value**
Number of cells (N)	3
Number of users per cell (K)	2
Number of antennas per BS (M)	8
Noise variance at individual user (σ_{ik}^2)	-127 dBm
The distance between two adjacent BSs	3 km
Array antenna gain (G_a)	15 dBi
Path loss model over a distance of ℓ km	$128.1 + 37.6\log_{10}(\ell)$
Angular offset standard deviation (σ_a)	2°
Log-normal shadowing standard deviation (σ_s)	10 dB
Per-BS transmit power restriction (P_i)	30 dBm

beamforming design in [11], the centralized worst-case robust power minimization design in [51] and the distributed worst-case robust power minimization design in [50] that takes no consideration of per-BS power restriction and assume bounded CSI uncertainties.

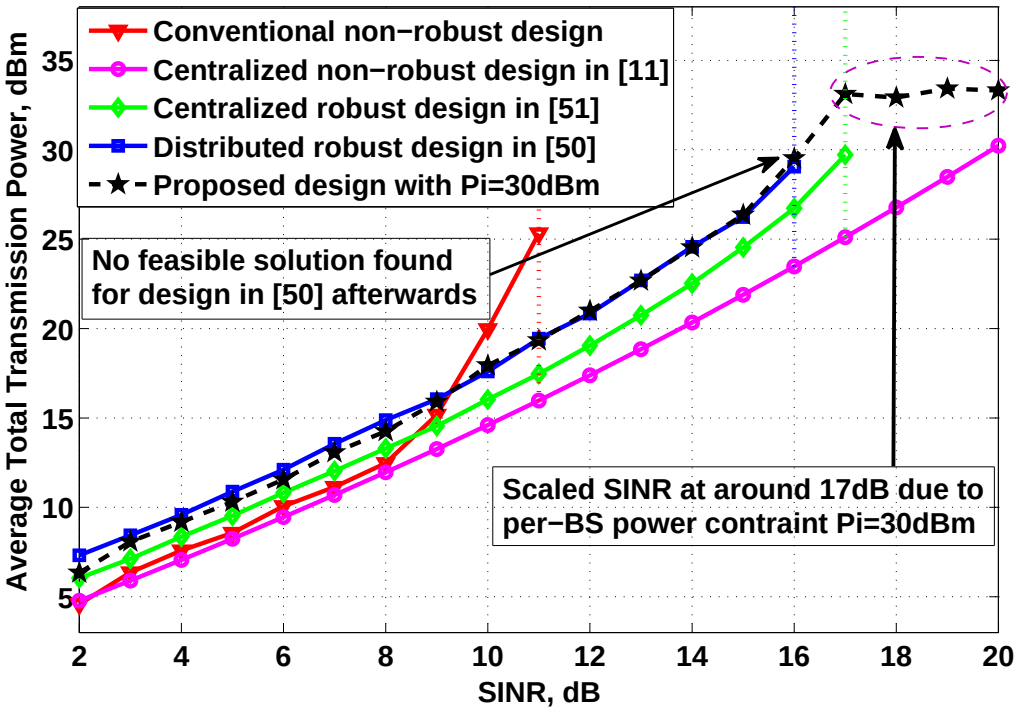


Figure 4.2: Comparison of total transmit power for different designs.

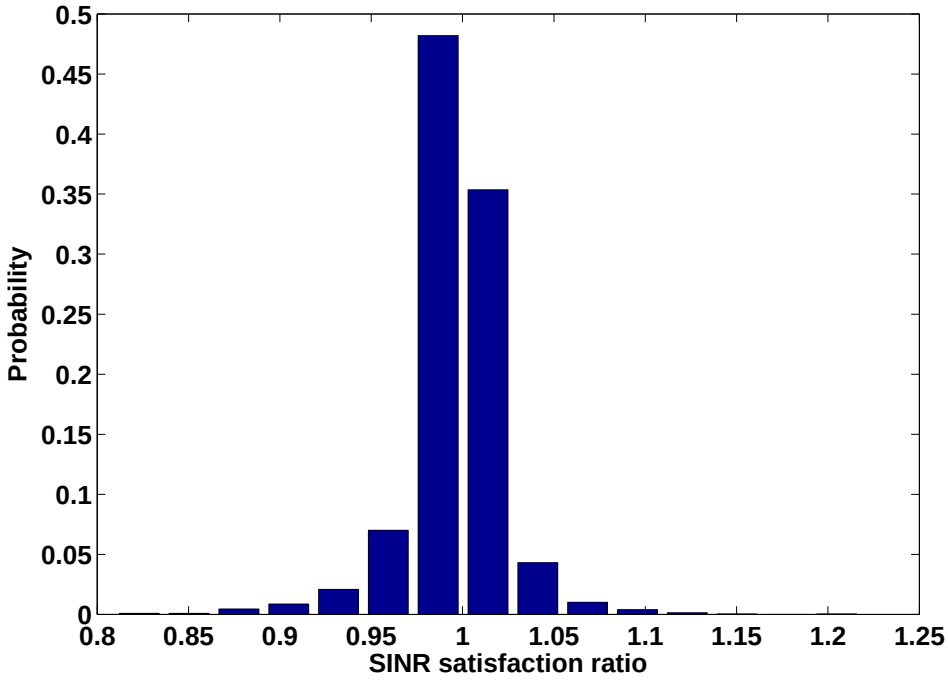
Fig. 4.2 presents the performance comparison of total transmit power for the proposed transmission strategy against other designs, under strict per-BS power constraint of 30 dBm. Note that the x-axis represents the target SINR γ_{ik}. As can be observed from the figure, the proposed strategy outperforms the conventional design in terms of expanding SINR operational range and closely follows its distributed robust counterpart in [50] until the per-BS power constraint is attained at around 16 dB of SINR target. When the SINR requirement is higher than 16 dB, the worst-case distributed design in [50] can not find a feasible solution due to the fact that it takes no consideration of individual BS transmit power constraints in their problem

formulation. Furthermore, no feasible beamforming solution can be provided by the worst-case centralized design in [51] for SINR requirements higher than 17 dB. On the contrary, although the per-BS power restriction limits the performance of the proposed strategy for high SINR requirements, it can provide a feasible solution at scaled desired SINR targets, with a total transmit power of $P_i = 30$ dBm. Thus, one may conclude that the proposed strategy is of practical significance, especially for dense users distribution since it optimally scales the SINR targets based on per BS power budgets and always provides a feasible solution at the scaled SINR target.

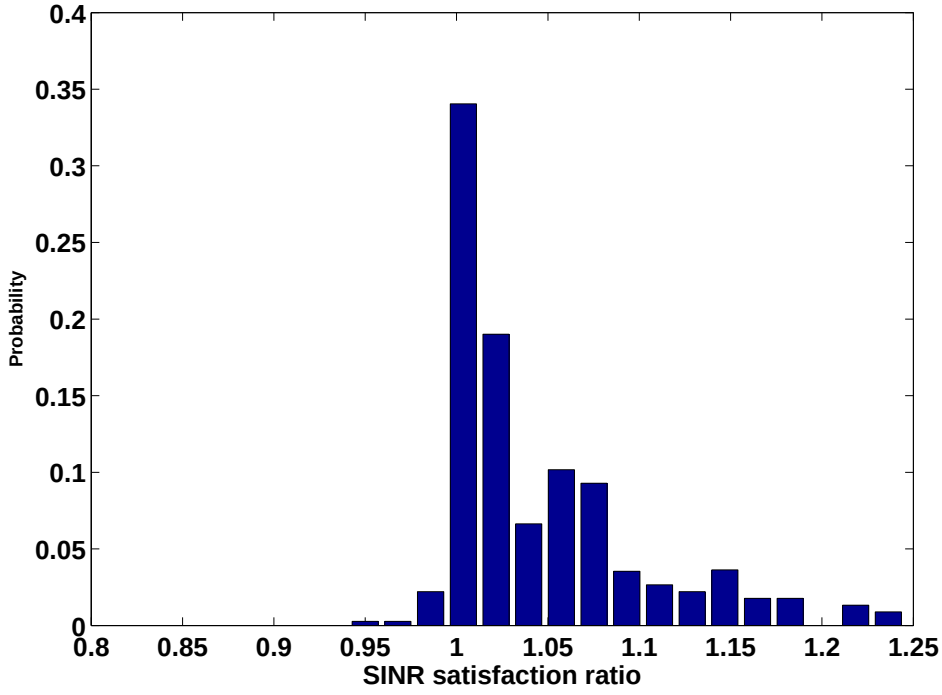
Let the SINR satisfaction ratio be defined as the achieved SINR over the scaled target SINR of UT_{ik}, i.e.,

$$\eta_{ik} = \frac{|\mathbf{h}_{iik}^H \mathbf{w}_{ik}|^2}{c_i \gamma_{ik} \left(\sum\limits_{\substack{n \neq k, \\ n \in \mathcal{L}_i}} |\mathbf{h}_{iik}^H \mathbf{w}_{in}|^2 + \sum\limits_{\substack{j \neq i, \\ j \in \mathcal{L}_b}} \sum\limits_{m \in \mathcal{L}_i} |\mathbf{h}_{jik}^H \mathbf{w}_{jm}|^2 + \sigma_{ik}^2 \right)}, \tag{4.24}$$

where $\eta_{ik} \geq 1$ indicates that the scaled SINR requirement of UT_{ik} is satisfied. Fig. 4.3 compares the average SINR satisfaction ratio at $\gamma = 10$ dB target SINR of the proposed decentralized robust transmission strategy against a non-robust power minimization design in [11] that assumes perfect knowledge of CSI. One can observe from the figure that for the proposed robust strategy that provides protection against channel uncertainties, almost all of the SINR satisfaction ratios stay above one. However, since the non-robust design in [11] provides no tolerance to any level of uncertainties, the actual achieved SINR fails to satisfy the SINR requirements for approximately 50 percent of the cases. Thus, one may conclude that the beamforming designs based on perfect CSI assumption may be sensitive to the channel uncertainties in a practical scenario. In comparison with Fig. 4.2, the performance gap between robust and non-robust designs can be interpreted as the cost for guaranteeing the worst-case quality of service at UTs, i.e., providing robustness against imperfect CSI.



(a)



(b)

Figure 4.3: Histograms of average SINR satisfaction ratio at $\gamma = 10$ dB of: a) non-robust power minimization design in [11], b) proposed robust strategy.

4.5 Concluding Remarks

This chapter studies a distributed robust approach for maximizing the weighted SINR targets at individual UTs in multi-cell interference networks. The problem is constrained to strict transmit power constraints at individual BSs in the presence of imperfect CSI. This problem is firstly mapped to an equivalent centralized aggregated transmit power minimization dual problem at individual BSs. Then the global-wise problem is decomposed into parallel subproblems via projected subgradient iterations to coordinate the ICI across BSs. Finally, a distributed UCB algorithm is proposed to find a global optimal trade-off between the weighted SINR targets and the per-BS transmit power constraints. Simulation results confirm the advantages of the proposed transmission strategy in providing larger SINR operational range and robustness against channel uncertainties in a multicell scenario with realistic parameter setup.

Chapter 5

A Bandit Approach to Price-Aware Energy Management in Cellular Networks

5.1 Introduction

Unlike Chapter 3 and Chapter 4 that focus on learning based joint intercell interference elimination and energy consumption optimization, this chapter mainly focuses on foresighted energy management that adapts to energy demand variations and contributes to the stable cost-efficient operation for green communication in future wireless communications networks. Accounting for the wireless channel random dynamism, a combinatorial multi-armed bandit (CMAB)-based reinforcement learning algorithm that benefits from an efficient exploration-exploitation trade-off is developed to minimize the time-averaged energy cost at individual base stations (BSs), powered by various energy markets and local renewable energy sources, over a finite time horizon. The proposed algorithm sustains traffic demands by enabling sparse beamforming to schedule dynamic user-to-BS allocation and proactive energy provisioning at BSs to make ahead-of-time price-aware energy management decisions.

5.1.1 Main Contribution

The main contributions of this chapter are summarized as follows.

- The proposed algorithm accounts for the inherent uncertain characteristics of the cellular communication networks by anticipating the amount of energy

demand ahead-of-time, purchasing it at a lower rate in the exploration mode and using this purchased energy in the following exploitation mode, so that the spot market energy provisioning at higher rate is minimized.

- The proposed algorithm enables smart scheduling that benefits from an efficient trade-off between the exploration (i.e., online training or learning) and the exploitation (i.e., operational) modes and reduces the exploration overhead. In addition, the two directional search in the exploration mode further improves the efficiency as compared with the single direction and full exploration learning algorithm proposed in [12].

Simulation results indicate a superior performance of the proposed algorithm in reducing the overall energy cost, as compared with with recently proposed non-learning based cooperative energy management designs in [4, 26] and a simplified CMAB based design in [12].

5.1.2 Organization

The rest of this chapter is organized as follows. Section 5.2 introduces the energy management model and downlink joint transmission model. In section 5.3, the cooperative energy management problem is formulated in a centralized manner and then transformed into numerically tractable form via semidefinite relaxation (SDR) technique and *reweighted* ℓ_1*-norm method*. Section 5.4 proposes an online learning algorithm inspired by CMAB model whilst the proposed strategy is analyzed and verified by the simulation results in section 5.5. Finally, section 5.6 concludes this chapter.

5.2 System Model

Consider a centralized cluster-based coordinated multipoint (CoMP) network in the downlink where a set of N BSs partially collaborate to serve K_i user terminals (UTs) over a shared bandwidth, as illustrated in Fig. 5.1. Each BS is equipped with

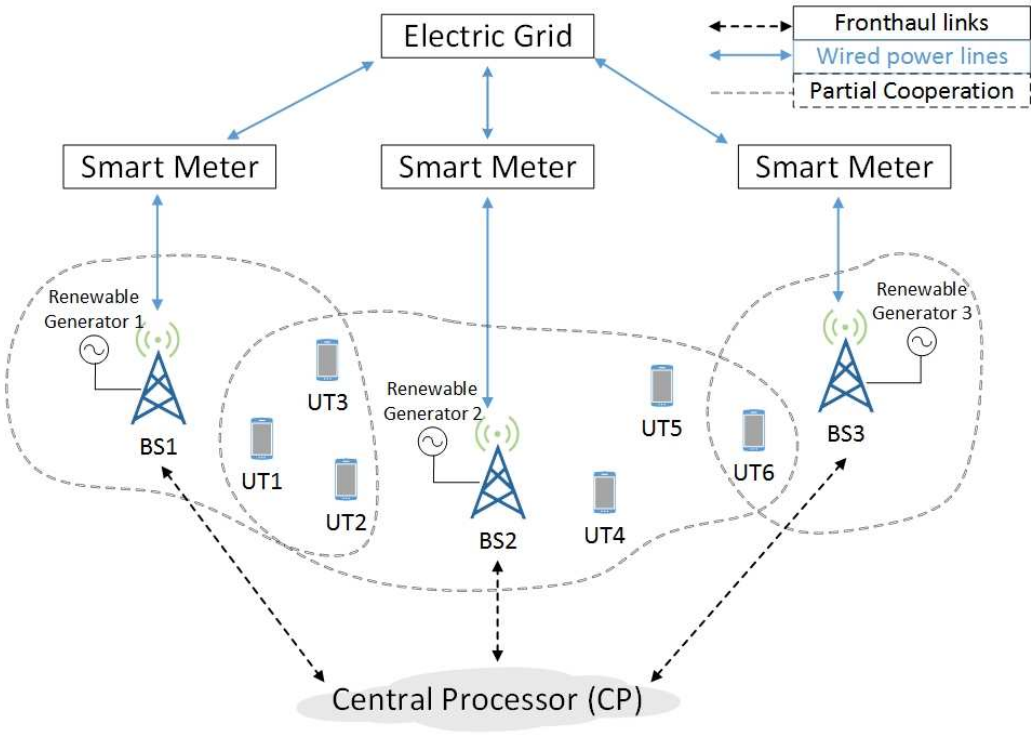


Figure 5.1: Illustration of downlink partial cooperation among BSs.

M antennas, whereas each UT has a single receiving antenna. Let $\mathcal{L}_b = \{1, \cdots, N\}$ and $\mathcal{L}_i = \{1, \cdots, K_i\}$ denote, respectively, the set of indexes of the BSs and the UTs within a cluster. The central processor (CP) coordinates all strategies based on perfect knowledge of channel state information and distributes all UTs' data to the corresponding BSs via finite-capacity fronthaul links. Besides, the CP also collects the energy information such as various energy market prices via the grid-deployed control links from the smart meters installed at individual BSs. The energy transmission between the electrical grid and the BSs is accomplished via dedicated power lines. Let the finite time horizon be divided into T discrete time slots indexed as $\mathcal{T} = \{1, \cdots, T\}$, such that the length of each time slot is smaller than the wireless channel coherence time. For convenience, the duration of a time slot is normalized to unity, thus the terms 'power' and 'energy' can be used interchangeably throughout this chapter. The proposed online learning algorithm in Section 5.4 runs over these time slots, such that an efficient trade-off between its exploration and exploitation modes

is achieved.

5.2.1 Energy Management Model

Assume no BS is equipped with frequently rechargeable energy storage device and the BSs are obliged to sell any excessive energy back to the grid. Let us also assume that at least one renewable energy generator is installed at the individual BS that can provide an amount of $G_n(t)$ units of renewable energy for the n-th BS at the t-th time slot, $t \in \mathcal{T}$, whilst BSs can access various energy markets at different prices. At the end of an exploration mode, an amount of $E_n^{[\mathrm{a}]}$ units of energy that can be sustained uniformly over a number of following exploitation time slots is purchased ahead-of-time for the n-th BS, $n \in \mathcal{L}_b$, at a price rate of $\pi^{[\mathrm{a}]}$. Let $E_n^{[\mathrm{a}]}(t)$ denote the ahead-of-time purchased energy allocated to the current time slot t. Let $E_n^{[\mathrm{r}]}(t)$ be the amount of real-time energy required to be purchased at time slot t due to both insufficient $E_n^{[\mathrm{a}]}(t)$ and the available renewable energy $G_n(t)$ at the n-th BS. Note that from the supply and demand perspective, $E_n^{[\mathrm{r}]}(t)$ in practice, should be purchased from the spot market at a higher price rate of $\pi^{[\mathrm{r}]}$, whereas $G_n(t)$ can be obtained locally at much lower rate of equivalent annual cost of renewable harvesters, i.e., $\pi^{[\mathrm{g}]}$. The surplus of available energy to a BS, i.e., $S_n(t)$, can be sold back to the grid at a fair rate of $\pi^{[\mathrm{e}]}$, i.e., $\pi^{[\mathrm{r}]} \geq \pi^{[\mathrm{a}]} \geq \pi^{[\mathrm{g}]} \geq \pi^{[\mathrm{e}]}$ [4]. The total energy cost incurred by the n-th BS at the t-th time slot can be written as [4]

$$C_n^{[\mathrm{total}]}(t) = \pi^{[\mathrm{r}]} E_n^{[\mathrm{r}]}(t) + \pi^{[\mathrm{a}]} E_n^{[\mathrm{a}]}(t) + \pi^{[\mathrm{g}]} G_n(t) - \pi^{[\mathrm{e}]} S_n(t). \tag{5.1}$$

Let $P_n^{[\mathrm{Tx}]}(t)$ and $P_n^{[\mathrm{c}]}$ be defined as the total transmit power from the n-th BS at the t-th time slot and the hardware circuit power consumption at the n-th BS, respectively. Then, the total energy consumption of the n-th BS at the t-th time slot, i.e., $P_n^{[\mathrm{total}]}(t)$, is upper-bounded by its energy budget [4, 5], i.e.,

$$P_n^{[\mathrm{total}]}(t) = \eta P_n^{[\mathrm{Tx}]}(t) + P_n^{[\mathrm{c}]} \leq G_n(t) + E_n^{[\mathrm{a}]}(t) + E_n^{[\mathrm{r}]}(t) - S_n(t), \tag{5.2}$$

where $\eta > 0$ denotes the power amplifier efficiency and $P_n^{[\mathrm{c}]}$ is assumed to be constant without loss of generality.

5.2.2 Downlink Transmission Model

Let $\mathbf{w}_{ni} \in \mathbb{C}^{M\times 1}$ and $\mathbf{h}_{ni} \in \mathbb{C}^{M\times 1}, n \in \mathcal{L}_b, i \in \mathcal{L}_i$ denote the beamforming vector and the channel vector from the n-th BS towards the i-th UT, respectively. Then, the signal received by the i-th UT can be expressed as the summation of the intended information-carrying signal of the i-th UT, the inter-user interference caused by all other non-desired information beams and the additive white Gaussian noise (AWGN) with variance of σ_i^2, i.e., $n_i \sim \mathbb{CN}(0, \sigma_i^2)$, as follows

$$z_i = \sum_{n\in\mathcal{L}_b} \mathbf{h}_{ni}^H \mathbf{w}_{ni} s_{ni} + \sum_{n\in\mathcal{L}_b} \sum_{j\neq i, j\in\mathcal{L}_i} \mathbf{h}_{ni}^H \mathbf{w}_{nj} s_{nj} + n_i. \tag{5.3}$$

Without loss of generality, let us assume that the transmitted symbols, i.e., s_{ni}, are independent and identically distributed and their transmission energy is normalized to one, i.e., $\mathbb{E}(s_{ni}) = 1$. The signal-to-interference-plus-noise ratio (SINR) at the i-th UT, $i \in \mathcal{L}_i$, is defined as

$$\mathrm{SINR}_i = \frac{|\sum\limits_{n\in\mathcal{L}_b} (\mathbf{h}_{ni}^H \mathbf{w}_{ni})|^2}{\sum\limits_{j\neq i, j\in\mathcal{L}_i} |\sum\limits_{n\in\mathcal{L}_b} (\mathbf{h}_{ni}^H \mathbf{w}_{nj})|^2 + \sigma_i^2}. \tag{5.4}$$

where σ_i^2 is the AWGN variance and is assumed to be identical at all UTs. The n-th BS's fronthaul capacity consumption, $n \in \mathcal{L}_b$, i.e., the summation of fronthaul data rate for transmitting data from the CP to the n-th BS, is given by [5]

$$B_n^{[\mathrm{fronthaul}]} = \sum_{i\in\mathcal{L}_i} \left\| \|\mathbf{w}_{ni}\|_2^2 \right\|_0 R_i, \quad \forall n \in \mathcal{L}_b, \tag{5.5}$$

where $R_i = \log_2(1+\mathrm{SINR}_i)$ is the achievable data rate (bit/s/Hz) for the i-th UT. The binary indicator function $\left\| \|\mathbf{w}_{ni}\|_2^2 \right\|_0$ that illustrates the scheduling choices between

the i-th UT and the n-th BS, is defined as

$$\left\|\|\mathbf{w}_{ni}\|_2^2\right\|_0 = \begin{cases} 0, & \text{if } \|\mathbf{w}_{ni}\|_2^2 = 0, \\ 1, & \text{if } \|\mathbf{w}_{ni}\|_2^2 \neq 0, \end{cases} \tag{5.6}$$

where $\|\mathbf{w}_{ni}\|_2^2 = 0$ implies that the i-th UT is not served by the n-th BS and, hence, the fronthaul link between the CP and the n-th BS is not used for coordinated transmission to the i-th UT.

5.3 Price-aware Energy Management

In accordance with (5.1), the total energy cost at the t-th time slot, $\forall t \in \mathcal{T}$, depends on a linear combination of the real-time trading variables, i.e., $E_n^{[\mathrm{r}]}(t)$ and $S_n(t)$, and the ahead-of-time energy purchase, i.e., $E_n^{[\mathrm{a}]}(t)$, given an available amount of renewable energy $G_n(t)$. We aim to minimize the total average energy cost over a finite time horizon via an online-learning assisted convex optimization. The downlink beamforming vectors and the real-time trading parameters, i.e., $E_n^{[\mathrm{r}]}(t)$ and $S_n(t)$, are the variables of the optimization problem. The ahead-of-time energy purchase $E_n^{[\mathrm{a}]}(t)$ is the learning parameter which is proactively determined by the proposed online learning strategy and fedback to the optimization problem. The convex optimization problem is formulated in the current Section and will then be integrated with the online learning strategy, introduced in Section 5.4, under Algorithm 5.4.2.

5.3.1 Problem Formulation

In order to minimize the energy cost at each time slot t, the optimization problem is formulated as

$$\begin{aligned}
\min_{\mathbf{w}_{ni}, E_n^{[\mathrm{r}]}(t), S_n(t)} \quad & \sum_{n\in\mathcal{L}_b} P_n^{[\mathrm{Tx}]}(t) + \sum_{n\in\mathcal{L}_b} \left\{E_n^{[\mathrm{r}]}(t)\right\} \qquad (5.7)\\
\text{s.t.} \quad & \mathrm{C1}: \mathrm{SINR}_i(t) \geq \gamma_i, \quad \forall i \in \mathcal{L}_i,\\
& \mathrm{C2}: B_n^{[\mathrm{fronthaul}]}(t) \leq B_n^{[\mathrm{limit}]}, \quad \forall n \in \mathcal{L}_b,\\
& \mathrm{C3}: \eta P_n^{[\mathrm{Tx}]}(t) + P_n^{[\mathrm{c}]} \leq G_n(t) + E_n^{[\mathrm{a}]}(t) - S_n(t) + E_n^{[\mathrm{r}]}(t), \ \forall n \in \mathcal{L}_b,\\
& \mathrm{C4}: P_n^{[\mathrm{Tx}]}(t) \leq P_n^{[\mathrm{Tmax}]}, \quad \forall n \in \mathcal{L}_b,\\
& \mathrm{C5}: E_n^{[\mathrm{r}]}(t) \geq 0, \quad \forall n \in \mathcal{L}_b,\\
& \mathrm{C6}: S_n(t) \geq 0, \quad \forall n \in \mathcal{L}_b,
\end{aligned}$$

where $P_n^{[\mathrm{Tx}]}(t) = \sum_{i\in\mathcal{L}_i} ||\mathbf{w}_{ni}||_2^2$ is the total transmit power of the n-th BS at the t-th time slot. C1 indicates the SINR constraint γ_i for the i-th UT and C2 represents the fronthaul link capacity restriction, i.e., $B_n^{[\mathrm{limit}]}$, for each BS. C3 emphasises that the individual BS's energy consumption is upper bounded by its energy budget, i.e., $G_n(t)$, $E_n^{[\mathrm{a}]}(t)$, $E_n^{[\mathrm{r}]}(t)$ and $S_n(t)$. C4 specifies the maximum transmit power, i.e., $P_n^{[\mathrm{Tmax}]}$, at the n-th BS. C5 and C6 indicate, respectively, that the spot market energy provisioning and the excessive energy to be sold back are non-negative.

5.3.2 Reweighted ℓ_1-norm and Semidefinite Programming

The optimization problem in (5.7) is NP-hard due to the non-convexity of the constraint C1 and the ℓ_0-norm term in C2. The intractable constraint C2 in (5.7) that formulates the sparse beamforming problem as ℓ_0-norm, is commonly handled with its ℓ_1-norm approximation via *reweighted ℓ_1-norm method* [104], as

$$B_n^{[\mathrm{fronthaul}]}(t) \approx \sum_{i\in\mathcal{L}_i} \left\| [\xi_{ni} \|\mathbf{w}_{ni}\|_2^2] \right\|_1 R_i = \sum_{i\in\mathcal{L}_i} \xi_{ni} \mathrm{tr}(\mathbf{w}_{ni}\mathbf{w}_{ni}^H) R_i. \qquad (5.8)$$

Algorithm 5.3.1. *Reweighted ℓ_1-norm method for solving problem in* (5.7)

1: **Initialize***: constant $\mu \to 0$, iteration count $s = 0$, weighting factor $\xi_{ni}(s) = 1$, maximum number of iterations s_{max}, $R_i(s) = log_2(1 + SINR_i)$.*

2: **While** *ξ_{ni} is not converged or $s \neq s_{max}$*

3: *Find the optimal beamforming vectors $\mathbf{w}^*_{ni}(s)$ by solving* (5.7)*;*

4: *Update the weighting factor $\xi_{ni}(s+1)$ as follows,*

$$\xi_{ni}(s+1) = \frac{1}{tr(\mathbf{w}^*_{ni}\mathbf{w}^{*H}_{ni})+\mu}, \quad \forall n \in \mathcal{L}_b, i \in \mathcal{L}_i;$$

5: *Calculate the achievable rate $R_i(s)$ as follows,*

$$R_i(s) = log_2[1 + \frac{tr(\sum\limits_{n\in\mathcal{L}_b} \mathbf{h}_{ni}\mathbf{h}^H_{ni}\mathbf{w}^*_{ni}(s)\mathbf{w}^{H*}_{ni}(s))}{\sum\limits_{j\in\mathcal{L}_i, j\neq i} tr(\sum\limits_{n\in\mathcal{L}_b} \mathbf{h}_{ni}\mathbf{h}^H_{ni}\mathbf{w}^*_{nj}(s)\mathbf{w}^{H*}_{nj}(s)) + \sigma_i^2}];$$

6: *Update $R_i(s+1) = R_i(s)$;*

7: *Increment the iteration number $s = s + 1$;*

8: **Endwhile**

In order to solve problem in (5.7) in time slot t, the cooperative links between the BSs and the UTs will be gradually and iteratively removed as per fronthaul link capacity constraints as well as the power budgets at the individual BSs, via alternating between solving optimal beamformer $\mathbf{w}^*_{ni}$ of problem (5.7) for a given weighting factor ξ_{ni}, and adjusting ξ_{ni} and R_i based on $\mathbf{w}^*_{ni}$, as detailed in Algorithm 5.3.1 [4, 104]. In particular, a BS transmitting with low transmit power to a particular UT in the s-th iteration will result in a large weighting factor $\xi_{ni}(s+1)$, which will lead to further reduction in the transmit power of that BS in the $(s+1)$-th iteration, until the convergence of ξ_{ni} is achieved. Once converged, the solution sparsity is attained, which is equivalent to turning off the BS for that particular UT, i.e., $\mathbf{w}^*_{ni} \approx 0$. It has been argued in [104] that the weighting factors could counteract the influence of the signal magnitude on the ℓ_1-norm surrogate to ℓ_0-norm, as ℓ_0-norm simply counts the number of nonzero elements of a vector and is not sensitive to their actual values.

Let us define $\mathbf{H}_{ni} = \mathbf{h}_{ni}\mathbf{h}^H_{ni}$ and semidefinite matrix $\mathbf{W}_{ni} = \mathbf{w}_{ni}\mathbf{w}^H_{ni}$. Then, the original problem in (5.7) can be transformed to a semidefinite programming (SDP)

problem after relaxing the rank-one constraints of rank $(\mathbf{W}_{ni}) = 1, \forall i \in \mathcal{L}_i, n \in \mathcal{L}_b$, as

$$\begin{aligned}
\min_{\mathbf{W}_{ni}, E_n^{[\mathrm{r}]}(t), S_n(t)} \quad & \sum_{n\in\mathcal{L}_b}\sum_{i\in\mathcal{L}_i}\mathrm{tr}(\mathbf{W}_{ni}) + \sum_{n\in\mathcal{L}_b}\left\{E_n^{[\mathrm{r}]}(t)\right\} \qquad (5.9)\\
\text{s.t.} \quad & \mathrm{C1}: \gamma_i^{-1}\mathrm{tr}(\sum_{n\in\mathcal{L}_b}\mathbf{H}_{ni}\mathbf{W}_{ni}) \geq \sum_{j\in\mathcal{L}_i, j\neq i}\mathrm{tr}(\sum_{n\in\mathcal{L}_b}\mathbf{H}_{ni}\mathbf{W}_{nj}) + \sigma_i^2, \forall i \in \mathcal{L}_i,\\
& \mathrm{C2}: \sum_{i\in\mathcal{L}_i}\xi_{ni}\mathrm{tr}(\mathbf{W}_{ni})R_i \leq B_n^{[\mathrm{limit}]}, \ \forall n \in \mathcal{L}_b,\\
& \mathrm{C3}: \eta\sum_{i\in\mathcal{L}_i}\mathrm{tr}(\mathbf{W}_{ni}) \leq G_n(t) + E_n^{[\mathrm{a}]}(t) - P_n^{[\mathrm{c}]} - S_n(t) + E_n^{[\mathrm{r}]}(t), \forall n \in \mathcal{L}_b,\\
& \mathrm{C4}: \sum_{i\in\mathcal{L}_i}\mathrm{tr}(\mathbf{W}_{ni}) \leq P_n^{[\mathrm{Tmax}]}, \forall n \in \mathcal{L}_b,\\
& \mathrm{C5}: E_n^{[\mathrm{r}]}(t) \geq 0, \quad \forall n \in \mathcal{L}_b,\\
& \mathrm{C6}: S_n(t) \geq 0, \quad \forall n \in \mathcal{L}_b,\\
& \mathrm{C7}: \mathbf{W}_{ni} \succeq 0, \quad \forall i \in \mathcal{L}_i, n \in \mathcal{L}_b.
\end{aligned}$$

Note that, if the obtained solutions $\mathbf{W}_{ni}^*$ are rank-one, the problem (5.9) yield same optimal solutions as problem (5.7).

Lemma 5.3.1. *The optimal solutions to the problems* (5.9) *satisfy rank* $(\mathbf{W}_{ni}^*) = 1$ *with probability one.*

Proof: Please refer to Appendix C [4].

5.4 Proactive Energy Management

Due to the combinatorial nature of distributed energy transmission from the grid to the BSs, the price-aware energy management problem studied in this chapter is classified as CMAB problem. The CMAB problem is defined as a system consists of J possible arms, where N arms, $N \subset J$, that form a super arm are played simultaneously and the reward of each arm is observed individually at each trial [84]. The objective is to maximize the long-term accumulated reward via a trade-off between observing the reward of new super arms, known as exploration (learning),

and proactively selecting the best-possible super arm for future time slots based on existing knowledge from the previous time slots, known as exploitation (operation).

In this chapter, each arm corresponds to a discrete ahead-of-time energy package to be selected for a BS and the reward of each arm corresponds to the difference between the energy cost at the t-th time slot and at the initial time slot. Thus, maximizing the accumulated reward is equivalent to minimizing the time-averaged energy cost. Let $\mathcal{K} = \{1, \cdots, K\}$ denote the set of indexes used to identify the learning (exploration) trials during a time slot, $\mathcal{J} = \{1, \cdots, J\}$ be the set of indexes associated to J arms, i.e., J ahead-of-time energy packages $\{\mathcal{E}^1, \cdots, \mathcal{E}^J\}$ offered by the grid, where $\mathcal{E}^e = \mathcal{E}^{e-1} + \Delta\mathcal{E}, e \in \mathcal{J}$. At the k-th trial, $k \in \mathcal{K}$, the CP selects a super arm, i.e., N ahead-of-time energy packages for N BSs, for next time slot, denoted by $\mathcal{S}^{[\text{set}]}(k) = \{E_1^{[\text{a}]}(k), \cdots, E_N^{[\text{a}]}(k)\}$. Let the individual reward of the arm $E_n^{[\text{a}]}(k)$ at the k-th trial be defined as [12]

$$\mathcal{R}(E_n^{[\text{a}]}(k)) = C_n^{[\text{total}]}(0) - C_n^{[\text{total}]}(k), \ \forall n \in \mathcal{L}_b, \tag{5.10}$$

where $C_n^{[\text{total}]}(0)$ and $C_n^{[\text{total}]}(k)$ are the total energy cost of the n-th BS at the initial trial of the initial time slot and the k-th trial of the current time slot, respectively, as per (5.1). Let $\mathbf{r_n^{[k,t]}} = (r_{n,1}^{[k,t]}, r_{n,2}^{[k,t]}, \cdots, r_{n,J}^{[k,t]})$ be defined as the reward vector of the n-th BS, where $r_{n,e}^{[k,t]}, e \in \mathcal{J}$, is the reward associated to the e-th ahead-of-time energy package in the k-th trial at the t-th time slot averaged over F independent channel realizations. Also let $\hat{\mathbf{r}}_\mathbf{n}^{[\mathbf{t}]} = (\hat{r}_{n,1}^{[t]}, \hat{r}_{n,2}^{[t]}, \ldots, \hat{r}_{n,J}^{[t]})$ and $\bar{\mathbf{r}}_\mathbf{n}^{[\mathbf{t}]} = (\bar{r}_{n,1}^{[t]}, \bar{r}_{n,2}^{[t]}, \ldots, \bar{r}_{n,J}^{[t]})$ denote mean reward vector and adjusted reward vector of individual ahead-of-time energy packages for the n-th BS at the t-th time slot, respectively.

In the sequel, an online learning algorithm to be executed at the CP, detailed in Fig. 5.2 as well as Algorithms 5.4.1 and 5.4.2, will be introduced to minimize the total energy cost over a finite time horizon. Similar to [89], the proposed algorithm enables smart scheduling that linearly increases the ratio of exploitation with an exponentially increased number of time slots, as presented in Fig. 5.3 and Table 5.1, which reduces the exploration overhead in terms of total energy cost over a finite time

horizon. The finite time horizon of T time slots is divided into P periods of increased length growing at a geometric progression, i.e., $T = 2(2^P - 1)$. Let $\mathcal{P} = \{1, \ldots, P\}$ denote the set of indexes of periods. In the p-th period that contains 2^p time slots, $p \in \mathcal{P}$, a total number of p time slots will be randomly selected as exploration mode whilst the rest time slots are reserved for exploitation mode. The principle of the smart scheduling is to reduce the fraction of time slots being selected as exploration mode with increasing period index, due to the fact that the estimation of the super arms' mean reward process is improved for a larger period index.

Table 5.1: Percentage of exploration using smart scheduling

Period index	1	2	3	4	5	6
No. of time slot	2	$2^2 = 4$	$2^3 = 8$	$2^4 = 16$	$2^5 = 32$	$2^6 = 60$
No. of exploration	1	2	3	4	5	6
% of exploration	0.50	0.50	0.429	0.333	0.242	0.167

In the exploration mode, Algorithm 5.4.1, i.e., *two directional super arm exploration at time slot t*, explores new super arm, i.e., new combination of ahead-of-time energy packages for N BSs, in a two directional way. More specifically, the exploring direction among all possible arms, i.e., forward or backward exploration, will be initially determined as described in step 9 and 11 of Algorithm 5.4.1, respectively, based on the rewards obtained at the current and the previous trials, followed by the super arm exploration for the next trial. The proposed Algorithm 5.4.1 guarantees that the individual BSs search in the proper direction towards the optimal arm that associated with the highest reward. Once a given number of K trials are completed, the mean reward for individual energy packages, i.e., $\hat{\mathbf{r}}_{\mathbf{n}}^{[\mathbf{t}]}$, for the n-th BS at the t-th time slot are estimated and adjusted within a controlled percentage, respectively, as per step 8 and 9 in Algorithm 5.4.2, i.e., *Online learning main algorithm*. The adjusted rewards, i.e., $\bar{\mathbf{r}}_{\mathbf{n}}^{[\mathbf{t}]}$, are first, averaged over all past time slots as per step 13, and then, used to update the index of optimal N arms, to be exploited in the next time slot, as detailed in step 14 of Algorithm 5.4.2. Note that by putting preference on the not frequently selected arms, the adjustment stage in

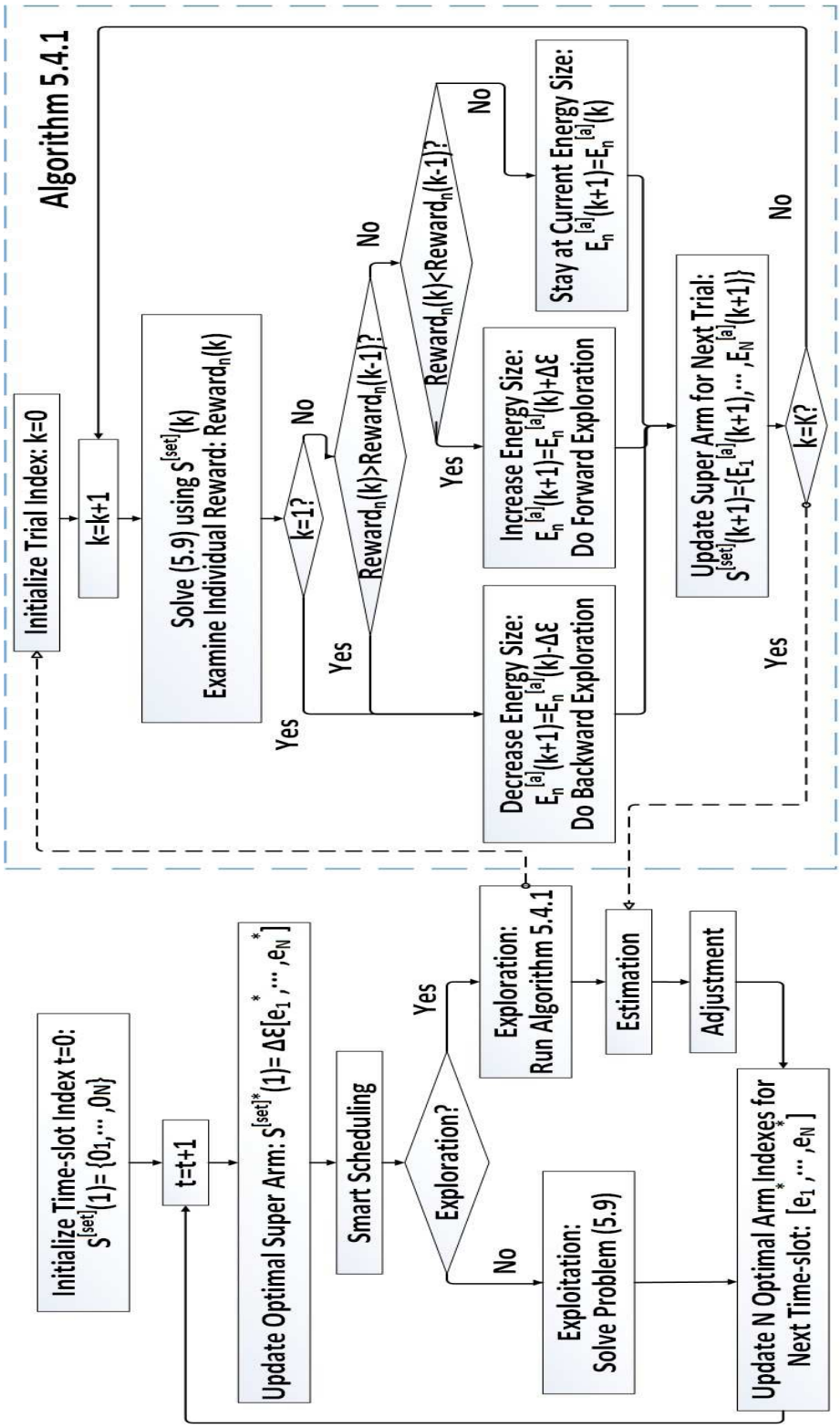


Figure 5.2: Flowchart diagram of proposed online learning algorithm

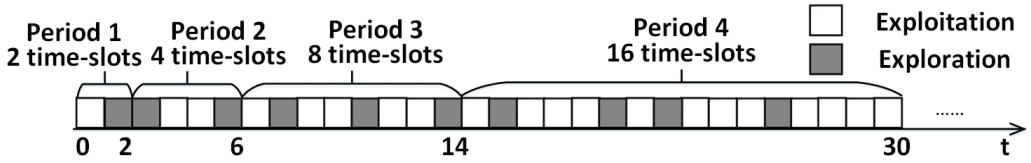


Figure 5.3: An exploration-exploitation trade-off model of smart scheduling

step 9 of Algorithm 5.4.2 encourages the CP to choose the least chosen arms as the starting points of future exploration time slots and examine the reward of those arms.

Algorithm 5.4.1. *Two Directional Super Arm Exploration at Time-slot t*

1: **For** $k = 1 : K$

2: *Solve problem in* (5.9),

3: *Compute* $C_n^{[total]}(k)$ *as per* (5.1) *and* $\mathcal{R}(E_n^{[a]}(k))$ *as per* (5.10),

4: **if** $k = 1$ *(initial trial) and* $E_n^{[a]}(k) \neq \mathcal{E}^1$

5: **then** $E_n^{[a]}(k+1) = E_n^{[a]}(k) - \Delta\mathcal{E}$,

6: **else if** $k = 1$ *(initial trial) and* $E_n^{[a]}(k) = \mathcal{E}^1$

7: **then** $E_n^{[a]}(k+1) = E_n^{[a]}(k) + \Delta\mathcal{E}$,

8: **else if** $\mathcal{R}(E_n^{[a]}(k)) > \mathcal{R}(E_n^{[a]}(k-1))$,

9: **then** *Do* **Backward Exploration**,

$$E_n^{[a]}(k+1) = E_n^{[a]}(k) - \Delta\mathcal{E},$$

10: **else if** $\mathcal{R}(E_n^{[a]}(k)) < \mathcal{R}(E_n^{[a]}(k-1))$,

11: **then** *Do* **Forward Exploration**,

$$E_n^{[a]}(k+1) = E_n^{[a]}(k) + \Delta\mathcal{E},$$

12: **else** $E_n^{[a]}(k+1) = E_n^{[a]}(k), \quad \forall n \in \mathcal{L}_b,$

13: **end if**

14: *Compute energy package index as* $e = \frac{E_n^{[a]}(k)}{\Delta\mathcal{E}},\ n \in \mathcal{L}_b,$

15: *Update* $r_{n,e}^{[k,t]} = \mathcal{R}(E_n^{[a]}(k)),\ \forall e \in \mathcal{J}, n \in \mathcal{L}_b,$

16: *Update* $\mathcal{S}^{[set]}(k+1) = \{E_1^{[a]}(k+1), \cdots, E_N^{[a]}(k+1)\}.$

17: **End for**

Algorithm 5.4.2. *Online Learning Main Algorithm*

1: **For** $t = 1 : T$

2: **if** $t = 1$ *(initial time slot)*

3: **then** *Initialize super arm as* $\mathcal{S}^{[set]}(1) = \{0_1, \cdots, 0_N\}$,

4: **else** *Update optimal super arm as*

$$\mathcal{S}^{[set]}(1)^* = \Delta\mathcal{E}[e_1^*, e_2^*, \cdots, e_N^*],$$

5: **end if**

6: **if** t *is selected for* **Exploration** *mode*

7: **then** *Run Algorithm 5.4.1,*

8: **Estimation Stage :**

Compute mean reward $\hat{\mathbf{r}}_\mathbf{n}^{[\mathbf{t}]} = (\hat{r}_{n,1}^{[t]}, \hat{r}_{n,2}^{[t]}, \ldots, \hat{r}_{n,\mathcal{J}}^{[t]})$,

where $\hat{r}_{n,e}^{[t]} = \frac{\sum_{k=1}^{K} r_{n,e}^{[k,t]}}{K},\ \forall e \in \mathcal{J}, n \in \mathcal{L}_b$,

9: **Adjustment Stage :**

Adjust $\bar{r}_{n,e}^{[t]} = \hat{r}_{n,e}^{[t]} + [\alpha\hat{r}_{n,e}^{[t]}, \sqrt{\frac{3lnt}{2\Psi_e}}]^-, \forall e \in \mathcal{J}, n \in \mathcal{L}_b$,

where α *is the step size and* Ψ_e *is number of times the e-th arm has been played,*

10: **else if** t *is selected for* **Exploitation** *mode*

11: *Solve problem in* (5.9),

12: **end if**

13: *Average* $\bar{\mathbf{r}}_\mathbf{n}^{[\mathbf{t}]}$ *over accumulated number of time slots, as*

$\bar{\mathbf{r}}_\mathbf{n} = \frac{\sum_{t'=1}^{t} \bar{\mathbf{r}}_\mathbf{n}^{[\mathbf{t}']}}{t} = [\bar{r}_{n,1}, \bar{r}_{n,2}, \cdots, \bar{r}_{n,J}], n \in \mathcal{L}_b$,

14: *For the next time slot: find* N *optimum arm indexes as*

$$e_n^* = \arg\max_e(\bar{r}_{n,e}), e \in \mathcal{J}, \forall n \in \mathcal{L}_b.$$

15: **End for**

5.5 Simulation Results

Consider a downlink system comprises 3 neighbouring 8-antennas BSs with a BS-BS distance of 500 m, transmitting toward 6 single-antenna UTs under a shared bandwidth, as shown in Fig. 5.4. A correlated channel model $\mathbf{h}_{ni} = \mathbf{C}_{ni}^{1/2}\mathbf{h}_w$ is adopted [1], where $\mathbf{h}_w \in \mathbb{C}^{M\times 1}$ are the zero-mean circularly

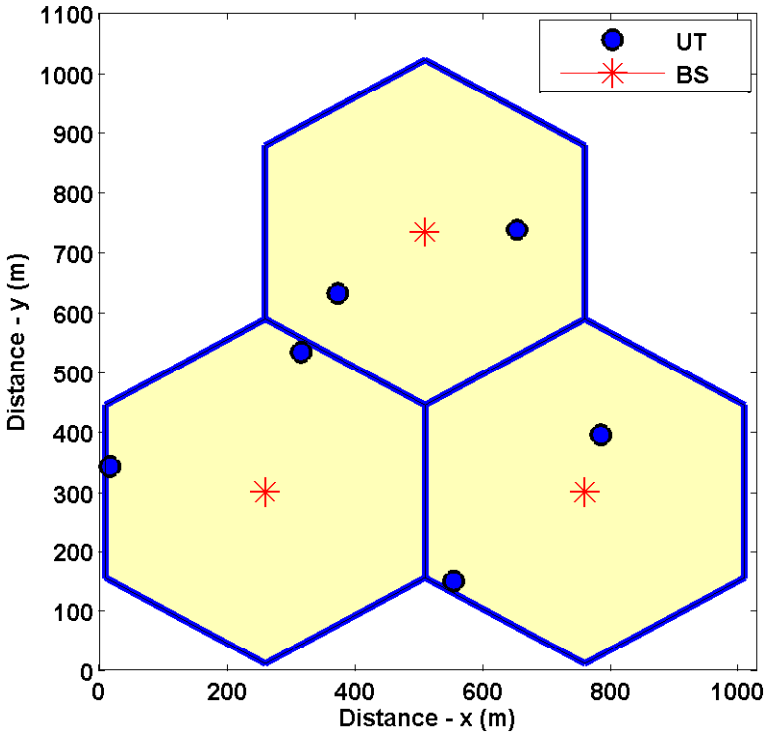


Figure 5.4: An example of multi-user downlink simulation topology.

symmetric complex Gaussian random variables with unit variance. $\mathbf{C}_{ni} \in \mathbb{C}^{M\times M}$ is the spatial covariance matrix with its (m,n)-th element given by $G_a L_p \sigma_F^2 e^{-0.5\frac{(\sigma_s \ln 10)^2}{100}} e^{j\frac{2\pi\delta}{\lambda}[(n-m)sin\theta]} e^{-2\left[\frac{\pi\delta\sigma_a}{\lambda}(n-m)cos\theta\right]^2}$ [1], where $G_a = 15$ dBi denotes the antenna gain, the path loss over a distance of ℓ km is modeled as $L_p(\text{dB}) = 125.2 + 36.3\log_{10}(\ell)$ [2], σ_F^2 is the variance of the complex Gaussian fading coefficient, $\sigma_s = 8$ dB is the log-normal shadowing standard deviation, $\sigma_a = 2^\circ$ is the angular offset standard deviation and θ is the estimated angle of departure. The renewable energy supplies at individual BSs at each time slot are, respectively, $G_1 = 1.5$ W, $G_2 = 0.2$ W and $G_3 = 0.05$ W, at a price of $\pi^{[\text{g}]} = £0.05$/W [4]. The noise figure at UTs and noise power spectral density are set to be 5 dB and -174 dBm/Hz, respectively. The simulation parameters are summarized in Table 5.2 [4–6]. The performance of the proposed strategy is evaluated with $K = 5$ learning trials averaging over $F = 20$ independent channel realizations for each time slot, for $T = 60$

Table 5.2: Simulation parameters [4–6]

Parameter	**Value**
Number of BSs (N)	3
Number of antennas per BS (M)	8
Number of the UTs (K_i)	6
Distance between two adjacent BSs	500 m
Renewable energy generation at BSs (G_1, G_2, G_3)	1.5 W, 0.2 W, 0.05 W
Per unit price of renewable energy ($\pi^{[\mathrm{g}]}$)	£0.05/W
Per unit price of ahead-of-time energy ($\pi^{[\mathrm{a}]}$)	£0.07/W
Per unit price of spot-market energy provisioning ($\pi^{[\mathrm{r}]}$)	£0.15/W
Per unit price of excessive energy sell ($\pi^{[\mathrm{e}]}$)	£0.02/W
Circuit power consumption at the n-th BS ($P_n^{[\mathrm{c}]}$)	30 dBm
Maximum transmit power allowance ($P_n^{[\mathrm{Tmax}]}$)	46 dBm
Fronthaul capacity limit at the n-th BS ($B_n^{[\mathrm{limit}]}$)	35 bits/s/Hz
Total number of time slots (T)	60
Total number of learning trials in each time slot (K)	5
The adjustment step size in Algorithm 5.4.2 (α)	0.5
Ahead-of-time energy packages offered at the grid	$\{100, 200, \cdots, 3000\}$ mW

time slots and $J = 30$ possible ahead-of-time energy packages with $\Delta\mathcal{E}$=100 mW, i.e., $\{\mathcal{E}^1, \mathcal{E}^2, \cdots, \mathcal{E}^J\} = \{100, 200, \cdots, 3000\}$ mW. The simulation results are obtained via CVX [35] using Intel i7-3770 CPU of 3.4GHz with 8GB RAM, and the running time for each learning trial is approximately 7 seconds without use of parallelization.

Fig. 5.5 and Fig. 5.6 compare the normalized total energy cost at $\gamma = 15$ dB target SINR of our proposed strategy against four designs, 1) a baseline joint energy trading and full cooperative energy management design in [26] that has no ahead-of-time energy purchase at all, 2) a non-learning based joint energy trading and partial cooperative energy management design in [4] that always purchases a fixed set of ahead-of-time energy packages, i.e., $E_1^{[\mathrm{a}]} = E_2^{[\mathrm{a}]} = E_3^{[\mathrm{a}]} = 700$ mW, 3) a simplified CMAB design in [12] that relaxes wireless channel dynamics and performs only single directional exploration mode without an efficient exploration-exploitation trade-off, and 4) the proposed strategy without smart scheduling. For fair comparison, identical constraints are applied to all designs. The total energy cost is normalized with respect to the initial value in the first time slot of the proposed strategy. In order to better

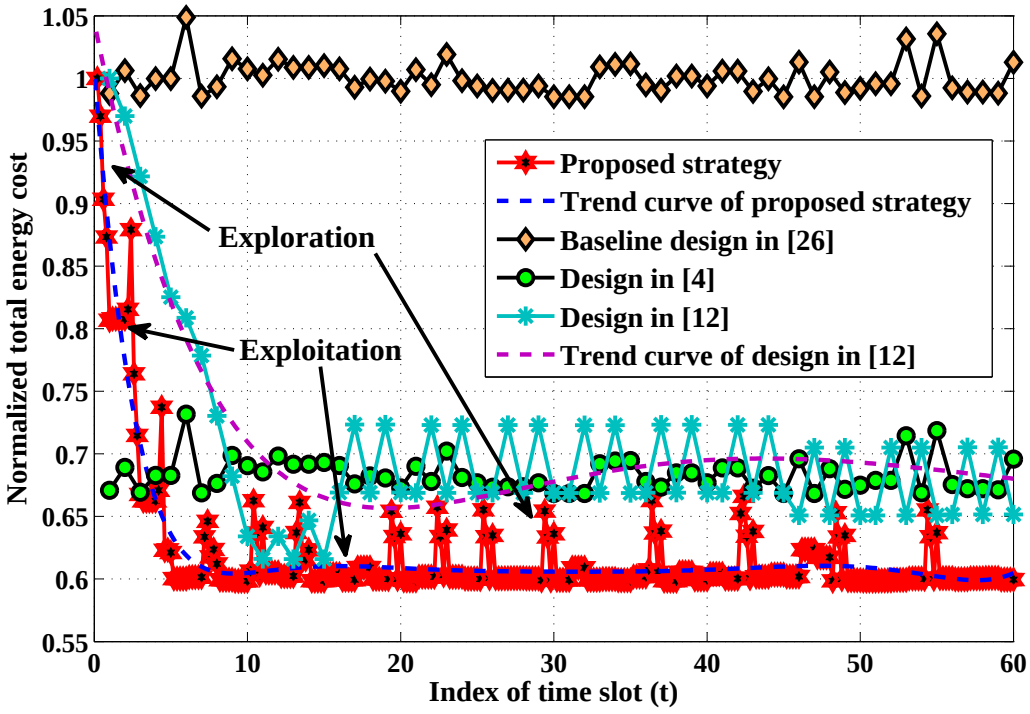


Figure 5.5: Normalized total energy cost of proposed strategy versus other designs at individual time slots at $\gamma = 15$ dB

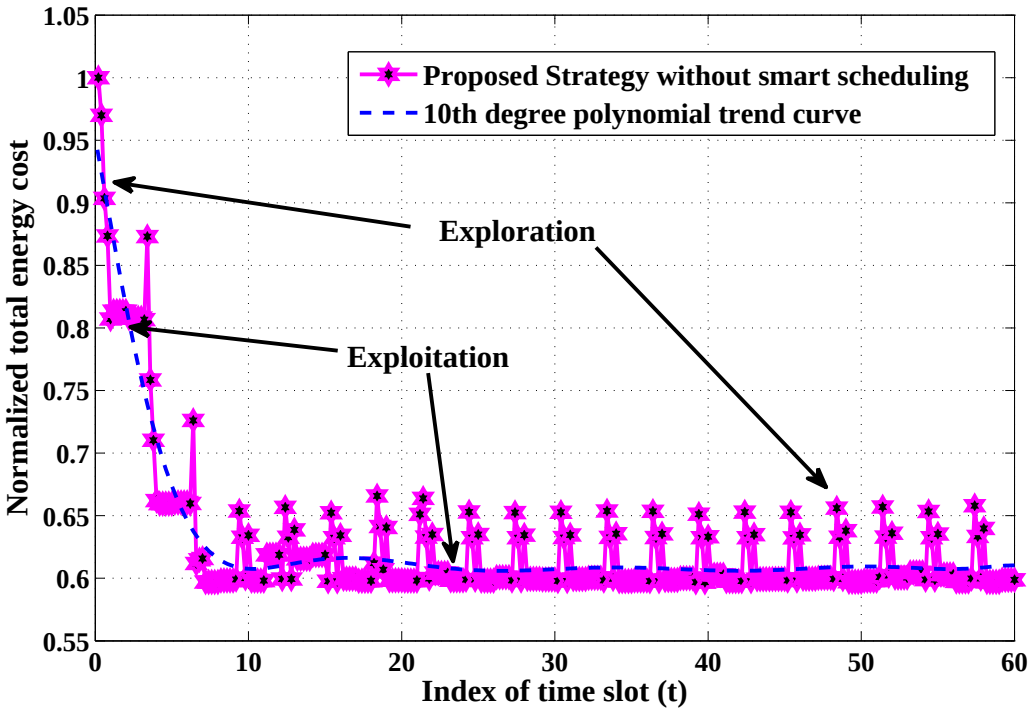


Figure 5.6: Normalized total energy cost of proposed strategy without smart scheduling at individual time slots at $\gamma = 15$ dB

evaluate the average performance and the convergence of the proposed strategy, the fitted 10th-degree-polynomial curve is adopted to represent the trend curve of the results.

The bursts and the smooth parts in Fig. 5.5, respectively, correspond to the exploration (online learning) mode and the exploitation (operational) mode. Note that the sharp jump in the beginning of an exploration is due to the adjustment stage via perturbation in step 9 of Algorithm 5.4.2 that prioritizes the least selected arms for the initial trial of exploration. It can be observed that the proposed Algorithm 5.4.2 guarantees the individual BSs searching in the right direction towards the optimal arm that associated with the highest reward. The fitted 10th-degree-polynomial trend curve to the results of the proposed strategy in Fig. 5.5 shows an improvement of approximately 40 percent over the initial state of the system from the 7th time-slot onwards. This is due to reducing significantly the real-time energy cost by ahead-of-time preparation for the future (i.e., real-time) energy demands at lower costs. Furthermore, an average percentage improvement of approximately 40, 8 and 7 per cent can be achieved by the proposed strategy as compared with [26], [4] and [12], respectively, due to the fact that their designs provide no adaption to the time-varying wireless channel conditions.

Recall from Section 5.4, the smart scheduling linearly increases the ratio of exploitation modes whilst decreases the proportion of high-energy-cost exploration modes with increasing number of time slots. The performance of the proposed strategy without smart scheduling is illustrated in Fig. 5.6, where a fixed trade-off between exploration and exploitation modes is adopted. The trend curve fitted to the results in Fig. 5.6 oscillates around the normalized energy cost of 0.61, as compared to 0.6 of the proposed strategy in Fig. 5.5. This difference in normalized energy cost is due to the fact that the proposed smart-scheduling-enabled strategy reduces the number of high-energy-cost exploration with increasing number of time slots as well as the better knowledge of the environment.

5.6 Concluding Remarks

This chapter proposes a CMAB approach to proactive price-aware energy management in cellular network, which adapts to dynamic wireless channel conditions and minimizes the overall energy cost over a finite time horizon. The proposed algorithm with smart scheduling reduces the exploration overhead by finding an efficient trade-off between exploring the rewards of new ahead-of-time energy purchase combinations, and exploiting the rewards of different combinations of ahead-of-time energy purchase acquired at the previous time slots. Simulation results confirm that in terms of cost-efficient energy provisioning at BSs, an average performance percentage improvement of 40, 8 and 7 per cent can be achieved by the proposed strategy as compared with recently proposed non-learning based designs in [26], [4] and a simplified CMAB based design in [12], respectively.

Chapter 6

Adaptive Energy Storage Management in Green Wireless Networks

6.1 Introduction

Unlike Chapter 5 that assumes no storage device is installed at individual base stations (BSs) and takes no consideration of the randomness of the renewable energy generation, this chapter mainly focuses on adaptive energy storage management in green wireless networks in the present of time-varying renewable energy supply. The dynamic nature of renewable energy generation not only introduces significant fluctuations on the electricity price, but can also destabilize the reliable and cost-efficient operation of the BSs supplied by hybrid grid and renewable energy generators. With the deployment of energy storage units at the demand side, the profit potentiality of the storage can be fully explored to compensate for not only the real-time energy shortage, but also the fluctuations of the electricity price, such that the long-term energy consumption cost can be minimized. Briefly, the challenge is how to integrate the randomness of the renewable energy generation with the main grid via predictive energy management for distributed energy storage devices at BSs.

6.1.1 Main Contribution

In order to address the dynamic statistics of wireless networks as well as the intermittent nature of renewable energy generation, this chapter develops an adaptive strategy inspired by combinatorial multi-armed bandit (CMAB) model for energy storage management and cost-aware coordinated load control at the BSs to minimize the average energy consumption cost at wireless networks in the long run.

This is a challenging task due to the following reasons. First, the state of each energy storage device is only known to the corresponding BS, but not the remaining BSs. Second, the actions of BSs are coupled in a complex way, which is unknown to them, and affect the overall energy cost. Third, the storage charging decisions have strong temporal correlations, i.e., the current decisions affect the future energy consumption costs, which induce temporal coupling in design variables. A novel adaptive algorithm is introduced to compensate for the randomness of the renewable energy generation via pre-charging the distributed storage devices. The proposed algorithm iteratively alternates between two decision making layers by exchanging conjectured information. The first layer located at the central processor (CP) designs the overall transmission strategy across the network of BSs using a convex semidefinite programming (SDP) and the second layer designs the pre-charging strategies for storages at distributed BSs via online learning, i.e., a CMAB approach.

Simulation results validate the superiority of the proposed strategy over a recently proposed storage-free learning-based design in [12].

6.1.2 Organization

The rest of this chapter is organized as follows. The system model is introduced in Section 6.2. In section 6.3, a cost-aware energy storage management problem is formulated in a centralized manner and transformed into numerically tractable form. Then, an adaptive storage management strategy inspired by CMAB model is proposed to minimize the time-averaged energy cost. Section 6.4 analyzes numerical simulation results and verifies the advantage of the proposed strategy against recent proposed designs. Finally, this chapter is summarized in Section 6.5.

6.2 System Model

Similar to Chapter 5, a downlink green wireless network is considered in this chapter, where a set of $\mathcal{L}_b = \{1, \cdots, N\}$ adjacent M-antenna BSs partially cooperated

to serve a set of $\mathcal{L}_i = \{1, \cdots, K_i\}$ single-antenna user terminals (UTs) over a shared bandwidth in accordance with their power budgets and fronthaul link capacity restrictions. Let us assume that the individual BSs are equipped with energy storage devices and are powered by local renewable energy generators, energy storage devices and the grid at various energy prices, as shown in Fig. 6.1. The storage-deployed BSs not only prevent the shortage of energy, but also enable the optimization of time-average energy cost via charging the storage either from the grid in advance at cheaper price or from the excessive renewable energy. Let the time horizon T be divided into discrete time slots, indexed as $\mathcal{T} = \{1, \cdots, T\}$, and assume the renewable energy supply varies across time slots but remains invariant within each time slot.

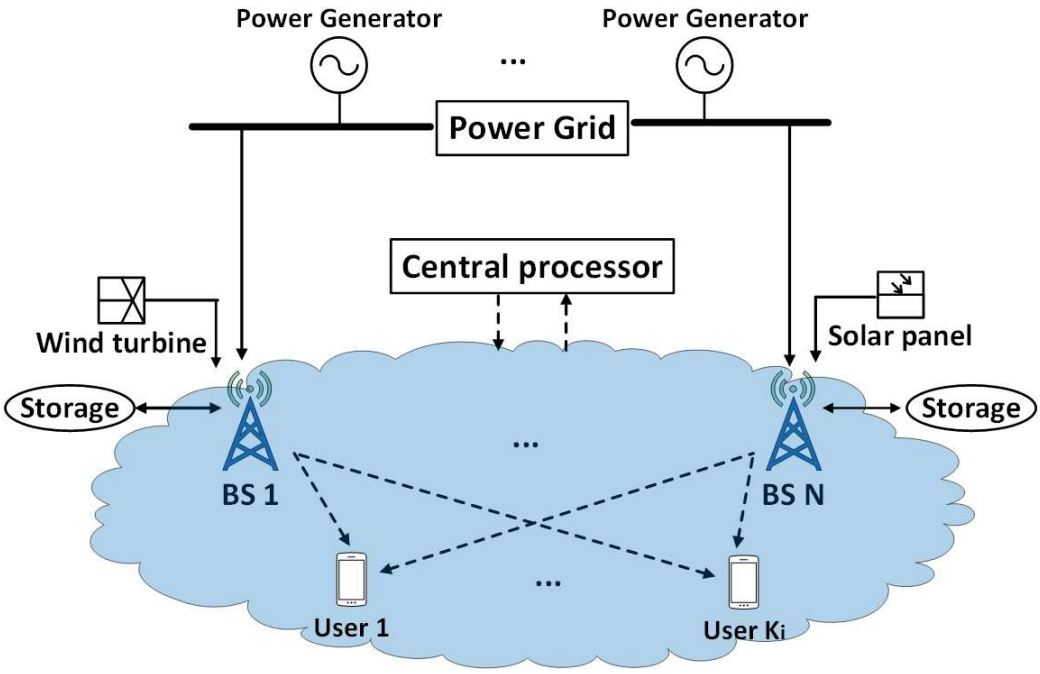


Figure 6.1: Illustration of downlink partial cooperation among storage-deployed BSs. The information flow is denoted by dashed lines and the energy flow is denoted by solid lines.

Similar to Section 5.2.1, let us assume that a varying amount of $G_n(t)$ units of renewable energy is generated at the n-th BS, $n \in \mathcal{L}_b$, at the t-th time slot, $t \in \mathcal{T}$. Let the amounts of $E_n^{[\mathrm{s}]}(t)$ and $E_n^{[\mathrm{c}]}(t)$ denote the units of the initial energy contents of the storage in the beginning of the t-th time slot and the units of energy charged

to the storage of the n-th BS prior to the actual time of energy demand at the t-th time slot, respectively. Notice that $E_n^{[\mathrm{s}]}(t) + E_n^{[\mathrm{c}]}(t) \in [0,\ E_n^{[\mathrm{capacity}]}]$, where $E_n^{[\mathrm{capacity}]}$ is the upper limit of the storage capacity at the n-th BS. Let an amount of $E_n^{[\mathrm{r}]}(t)$ units of energy be the energy shortage to be real-time supplied by the grid to the n-th BS at the t-th time slot. Let $\pi^{[\mathrm{r}]}, \pi^{[\mathrm{c}]}, \pi^{[\mathrm{g}]}$ and $\pi^{[\mathrm{s}]}$ be denoted, respectively, as the per unit energy prices for $E_n^{[\mathrm{r}]}(t)$, $E_n^{[\mathrm{c}]}(t)$, the per unit equivalent annual cost of renewable harvesters for $G_n(t)$ and the per unit equivalent annual cost of storage devices for storing an amount of $E_n^{[\mathrm{s}]}(t)$ units of energy, respectively. Similar to [27], it is assumed that $\pi^{[\mathrm{r}]} \geq \pi^{[\mathrm{c}]} \geq \pi^{[\mathrm{g}]} \geq \pi^{[\mathrm{s}]}$, such that the storage device will be charged when necessary and the renewable energy generation can be fully utilized. Then, the total energy cost of the n-th BS at the t-th time slot, i.e., $C_n^{[\mathrm{total}]}(t)$, is given by [4]

$$C_n^{[\mathrm{total}]}(t) = \pi^{[\mathrm{r}]} E_n^{[\mathrm{r}]}(t) + \pi^{[\mathrm{c}]} E_n^{[\mathrm{c}]}(t) + \pi^{[\mathrm{g}]} G_n(t) + \pi^{[\mathrm{s}]} E_n^{[\mathrm{s}]}(t). \tag{6.1}$$

Similar to Section 5.2.1, the total energy consumption of the n-th BS at the t-th time slot is upper-bounded by the energy budget at the n-th BS [26], as

$$\eta P_n^{[\mathrm{Tx}]}(t) + P_n^{[\mathrm{c}]} \leq G_n(t) + E_n^{[\mathrm{s}]}(t) + E_n^{[\mathrm{c}]}(t) + E_n^{[\mathrm{r}]}(t), \tag{6.2}$$

It is clear that at the end of the t-th time slot, the surplus energy that can be charged to the storage of the n-th BS is $[G_n(t) + E_n^{[\mathrm{s}]}(t) + E_n^{[\mathrm{c}]}(t) + E_n^{[\mathrm{r}]}(t) - P_n^{[\mathrm{Tx}]}(t) - P_n^{[\mathrm{c}]}]^+$. Thus, the initial energy storage of the n-th BS at the $(t+1)$-th time slot is constrained by the following expression [7]:

$$\begin{aligned} E_n^{[\mathrm{s}]}(t+1) = \min\{E_n^{[\mathrm{capacity}]},\quad & \max\{G_n(t) + E_n^{[\mathrm{s}]}(t) + E_n^{[\mathrm{c}]}(t) \\ & + E_n^{[\mathrm{r}]}(t) - P_n^{[\mathrm{Tx}]}(t) - P_n^{[\mathrm{c}]},\ 0\}\}. \end{aligned} \tag{6.3}$$

6.3 Adaptive Storage Management Strategy

In the sequel, an adaptive storage management algorithm used jointly by the CP to iteratively update the downlink beamforming vectors, i.e., $\mathbf{w}_{ni}, n \in \mathcal{L}_b, i \in \mathcal{L}_i$, at BSs as well as the amount of real-time energy supplied by the grid, i.e., $E_n^{[\mathrm{r}]}(t), t \in \mathcal{T}$, and by the individual BSs to update their strategies of charging their locally installed storage devices , i.e., $E_n^{[\mathrm{c}]}(t), t \in \mathcal{T}$, will be introduced in order to efficiently compensate for the randomness of the renewable generations. Individual BSs send their conjectured amount of required storage charges $E_n^{[\mathrm{c}]}(t)$ to the CP and receive the corresponding instantaneous reward from the CP. This process of iterative exchange of data allows the proposed adaptive algorithm to converge to optimal conjectured optimization variables, i.e., $\mathbf{w}_{ni}^*$, $E_n^{[\mathrm{r}]}(t)^*$ and the amount of energy charge to be deposited to the storage devices at a current time slot $E_n^{[\mathrm{c}]}(t)$.

6.3.1 Problem Formulation

Due to the combinatorial nature of distributed deployment of the energy storage devices across the BSs, the problem of adaptive storage energy management is formulated as a reinforcement learning problem based on CMAB model that is governed by a trade-off between exploring new sets of arms and exploiting the best set of arms to insure the time-averaged cost efficiency of the BSs over a time horizon "T", as illustrated in Fig. 6.2. Let us consider a set of arms denoted as a super arm, where each arm corresponds to an energy size to be stored in the storage of a BS in advance of the actual time that the shortage of energy may occur. A super arm is comprised of N arms chosen for N BSs out of J possible arms, i.e., $N \subset J$. Let us define the reward of the arm chosen for the n-th BS at time slot t, as

$$\mathcal{R}_n(t) = C_n^{[\mathrm{total}]}(0) - C_n^{[\mathrm{total}]}(t), \ \forall n \in \mathcal{L}_b, t \in \mathcal{T}, \tag{6.4}$$

where $C_n^{[\mathrm{total}]}(0)$ and $C_n^{[\mathrm{total}]}(t)$ are the total energy cost of the n-th BS at the initial time slot and at the t-th time slot, respectively. The proposed CMAB based adaptive

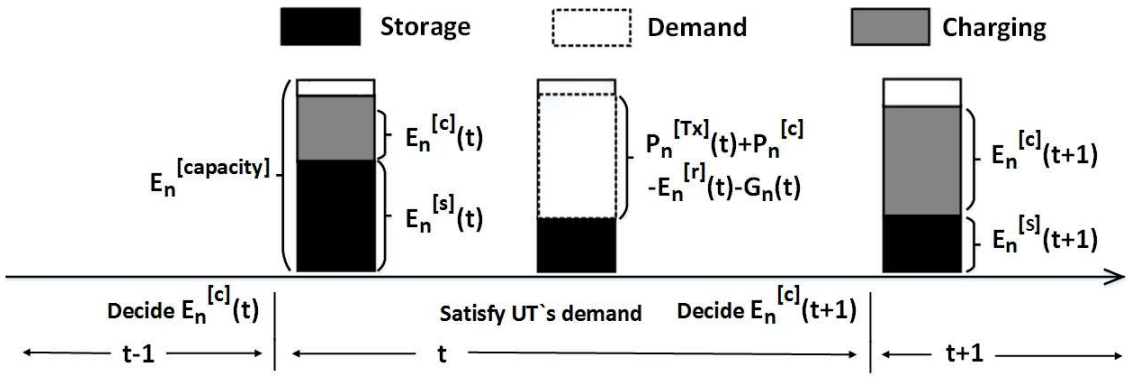


Figure 6.2: Illustration of proposed energy storage management strategy

algorithm maximizes the time-averaged accumulated reward over the online decisions on the amount of electricity to be stored in the storage devices of individual BSs, as

$$\max_{E_n^{[c]}(t)} \left\{ \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n\in\mathcal{L}_b} \mathcal{R}_n(t) \right\}. \tag{6.5}$$

Similar to Section 5.3.1, the energy consumption at individual BSs at time slot t is governed by the following optimization problem for resource allocation.

$$\begin{aligned} &\min_{\mathbf{w}_{ni}, E_n^{[\mathrm{r}]}(t)} \quad \sum_{n\in\mathcal{L}_b} P_n^{[\mathrm{Tx}]}(t) + \max_{n\in\mathcal{L}_b} \left\{ E_n^{[\mathrm{r}]}(t) \right\} \qquad (6.6)\\ \text{s.t.} \quad &\mathrm{C1}: \mathrm{SINR}_i(t) \geq \gamma_i, \quad \forall i \in \mathcal{L}_i,\\ &\mathrm{C2}: B_n^{[\mathrm{fronthaul}]}(t) \leq B_n^{[\mathrm{limit}]}, \quad \forall n \in \mathcal{L}_b,\\ &\mathrm{C3}: \eta P_n^{[\mathrm{Tx}]}(t) + P_n^{[\mathrm{c}]} - G_n(t) - E_n^{[\mathrm{s}]}(t) - E_n^{[\mathrm{c}]}(t) \leq E_n^{[\mathrm{r}]}(t), \quad \forall n \in \mathcal{L}_b,\\ &\mathrm{C4}: P_n^{[\mathrm{Tx}]}(t) \leq P_n^{[\mathrm{Tmax}]}, \ \forall n \in \mathcal{L}_b,\\ &\mathrm{C5}: E_n^{[\mathrm{r}]}(t) \geq 0, \ \forall n \in \mathcal{L}_b, \end{aligned}$$

where C3 indicates that the energy shortage of the n-th BS will be provisioned by the grid as per (6.2), whilst $E_n^{[\mathrm{s}]}(t)$ is updated as

$$\begin{aligned} E_n^{[\mathrm{s}]}(t) = \min\{E_n^{[\mathrm{capacity}]}, \quad &\max\{G_n(t-1) - P_n^{[\mathrm{c}]} - P_n^{[\mathrm{Tx}]}(t-1) \qquad (6.7)\\ &+E_n^{[\mathrm{s}]}(t-1) + E_n^{[\mathrm{c}]}(t-1) + E_n^{[\mathrm{r}]}(t-1), \ 0\}\}, \end{aligned}$$

in the beginning of the t-th time slot.

6.3.2 SDP Optimization

Let us define $\mathbf{W}_{ni} = \mathbf{w}_{ni}\mathbf{w}_{ni}^H$ and $\mathbf{H}_{ni} = \mathbf{h}_{ni}\mathbf{h}_{ni}^H$. By inclusion of the online learning process to decouple the time coupled constraints, the original problem in (6.6) can be simplified to an SDP optimization problem at the t-th time slot after adopting the *reweighted ℓ_1-norm method* in Section 5.3.2 and relaxing the rank-one constraints of $\text{rank}(\mathbf{W}_{ni}) = 1$, as

$$\begin{aligned}
&\min_{\mathbf{W}_{ni},\chi} \quad \sum_{n\in\mathcal{L}_b}\sum_{i\in\mathcal{L}_i} \text{tr}(\mathbf{W}_{ni}) + \chi \qquad (6.8)\\
\text{s.t.} \quad &\text{C1}: \gamma_i^{-1}\text{tr}(\sum_{n\in\mathcal{L}_b} \mathbf{H}_{ni}\mathbf{W}_{ni}) \geq \sum_{j\in\mathcal{L}_i, j\neq i} \text{tr}(\sum_{n\in\mathcal{L}_b} \mathbf{H}_{ni}\mathbf{W}_{nj}) + \sigma_i^2,\ \forall i \in \mathcal{L}_i,\\
&\text{C2}: \sum_{i\in\mathcal{L}_i} \xi_{ni}\text{tr}(\mathbf{W}_{ni})R_i \leq B_n^{[\text{limit}]}, \quad \forall n \in \mathcal{L}_b,\\
&\text{C3}: \eta\sum_{i\in\mathcal{L}_i} \text{tr}(\mathbf{W}_{ni}) + P_n^{[\text{c}]} - E_n^{[\text{s}]}(t) - E_n^{[\text{c}]}(t) - G_n(t) \leq E_n^{[\text{r}]}(t),\ \forall n \in \mathcal{L}_b,\\
&\text{C4}: \sum_{i\in\mathcal{L}_i} \text{tr}(\mathbf{W}_{ni}) \leq P_n^{[\text{Tmax}]}, \quad \forall n \in \mathcal{L}_b,\\
&\text{C5}: E_n^{[\text{r}]}(t) \geq 0, \quad \forall n \in \mathcal{L}_b,\\
&\text{C6}: E_n^{[\text{r}]}(t) \leq \chi,\ \forall n \in \mathcal{L}_b.\\
&\text{C7}: \mathbf{W}_{ni} \succeq 0, \quad \forall i \in \mathcal{L}_i, n \in \mathcal{L}_b,
\end{aligned}$$

Lemma 6.3.1. *The optimal solutions to the problems* (6.8) *satisfy rank* $(\mathbf{W}_{ni}^*) = 1$ *with probability one.*

Proof: Please refer to a similar proof as in Appendix C.

6.3.3 Proposed Online Learning Algorithm

This section introduces a CMAB-inspired online learning algorithm, detailed in Algorithm 6.3.1, to guarantee BSs' cost efficient operation in the long run. The purpose of the online learning part of the proposed algorithm at individual BSs is

to determine proactively the optimal conjectured amount of storage charging, i.e., $E_n^{[\mathrm{c}]}(t)$, ahead of time, before experiencing a possible energy shortage at the time slot "t", such that when the CP reacts based on that, the resulting transmission strategy, i.e., the beamforming vectors $\{\mathbf{w}_{ni}(t)\}_{t,i}$ and the supporting real-time amount of energy supply from the grid $\{E_n^{[\mathrm{r}]}(t)\}_t$, minimizes the overall energy cost of the network.

Similar to Section 5.4, let $\mathcal{K} = \{1, \cdots, K\}$, $\mathcal{J} = \{1, \cdots, J\}$ and $\mathcal{S}^{[\mathrm{set}]}(k) = \{E_1^{[\mathrm{c}]}(k), \cdots, E_N^{[\mathrm{c}]}(k)\}$ denote, respectively, the set of indexes of the learning trials during an exploration time slot, the set of indexes associated to J arms, i.e., J discrete energy charging sizes $\{\mathcal{E}^1, \cdots, \mathcal{E}^J\}$ with difference of $\Delta\mathcal{E}$, and the selected super arm that consists of N energy sizes to be stored at N BSs' storage devices in the k-th learning trial of a time slot, $k \in \mathcal{K}$. Let us define the reward of the arm selected for the n-th BS at the k-th learning trial in the t-th time slot as

$$\mathcal{R}_t(E_n^{[\mathrm{c}]}(k)) = C_n^{[\mathrm{total}]}(0) - C_n^{[\mathrm{total}]}(k), \ \forall n \in \mathcal{L}_b, k \in \mathcal{K}, t \in \mathcal{T}, \tag{6.9}$$

During time slots allocated for exploration, individual new super arms are explored and the set of energy charging sizes to the BSs' storage devices is assigned for the next learning trial based on the rewards acquired from the current and the previous learning trials. Then, the mean rewards for individual arms assigned to the n-th BS's storage device during the t-th time slot, i.e., $\hat{\mathbf{r}}_{\mathbf{n}}^{[\mathbf{t}]}$, are estimated and adjusted as per the steps 22 and 23 of Algorithm 6.3.1, respectively. The adjustment step 23 in Algorithm 6.3.1 implements the trade-off between exploiting the set of arms resulted in the highest accumulated reward so far and exploring new sets of arms that are not frequently selected and may result in a better accumulated reward during the future time slots. The proposed algorithm by design is not sensitive to the time scale due to the fact that the exploration cycle of Algorithm 6.3.1 responds to the variation in the environment by making adaptive decisions of $E_n^{[\mathrm{c}]}(t)$ for the upcoming exploitation cycles based on long-term time averaged accumulated rewards with a discount factor of $\mathcal{D}$ that indicates the importance of previous rewards, as detailed in step 27 in

Algorithm 6.3.1.

Algorithm 6.3.1. *Adaptive storage management algorithm*

For $t = 1 : T$

if $t = 1$ *(initial time slot)*

then *Initialize* $E_n^{[s]}(1) = 0$, $E_n^{[capacity]}$, *and* $\mathcal{S}^{[set]}(1) = \{0_1, \cdots, 0_N\}$,

else $\mathcal{S}^{[set]*}(1) = \Delta\mathcal{E}[e_1^*, e_2^*, \cdots, e_N^*]$,

end if

if t *is* **Exploration**

For $k = 1 : K$

Solve problem in (6.8),

Compute $C_n^{[total]}(k)$ *as per* (6.1) *and* $\mathcal{R}_t(E_n^{[c]}(k))$ *as per* (6.9),

if $k = 1$ *(initial trial) and* $E_n^{[c]}(k) \neq \mathcal{E}^J$

then $E_n^{[c]}(k+1) = E_n^{[c]}(k) + \Delta\mathcal{E},\ n \in \mathcal{L}_b$,

else if $\mathcal{R}_t(E_n^{[c]}(k)) < \mathcal{R}_t(E_n^{[c]}(k-1))$,

then *Backward search as* $E_n^{[c]}(k+1) = E_n^{[c]}(k) - \Delta\mathcal{E}$,

else if $\mathcal{R}_t(E_n^{[c]}(k)) > \mathcal{R}_t(E_n^{[c]}(k-1))$

then *Forward search as* $E_n^{[c]}(k+1) = E_n^{[c]}(k) + \Delta\mathcal{E}$,

else $E_n^{[c]}(k+1) = E_n^{[c]}(k)$,

end if

Compute the arm index e *as* $e = \frac{E_n^{[c]}(k)}{\Delta\mathcal{E}}, n \in \mathcal{L}_b$,

Update the reward vector of the n*-th BS in the* k*-th trial, i.e.,*
$\mathbf{r_n^{[k,t]}} = (r_{n,1}^{[k,t]}, r_{n,2}^{[k,t]}, \cdots, r_{n,J}^{[k,t]})$, *as* $r_{n,e}^{[k,t]} = \mathcal{R}_t(E_n^{[c]}(k)),\ \forall e \in \mathcal{J}, n \in \mathcal{L}_b$,

Update super arm for next trial as
$\mathcal{S}^{[set]}(k+1) = \{E_1^{[c]}(k+1), \cdots, E_N^{[c]}(k+1)\}$.

End for

Estimation Stage :

Compute the estimated mean reward vector, i.e., $\hat{\mathbf{r}}_\mathbf{n}^{[\mathbf{t}]} = (\hat{r}_{n,1}^{[t]}, \hat{r}_{n,2}^{[t]}, \ldots, \hat{r}_{n,J}^{[t]})$,
as $\hat{r}_{n,e}^{[t]} = \frac{\sum_{k=1}^{K} r_{n,e}^{[k,t]}}{K},\ \forall e \in \mathcal{J}$,

Adjustment Stage :

Table 6.1: Simulation parameters [4–7]

Parameter	**Value**
Number of BSs (N)	3
Number of antennas per BS (M)	8
Number of the UTs (K_i)	6
Distance between two adjacent BSs	500 m
Renewable energy generation at BS 1 (G_1)	[0.5 1.0] W
Renewable energy generation at BS 2 (G_2)	[0.1 0.5] W
Renewable energy generation at BS 3 (G_3)	[0.03 0.1] W
Per unit price of renewable energy ($\pi^{[\mathrm{g}]}$)	£0.05/W
Per unit price of ahead-of-time battery charging ($\pi^{[\mathrm{c}]}$)	£0.07/W
Per unit price of real-time energy provisioning ($\pi^{[\mathrm{r}]}$)	£0.15/W
Per unit price for storing energy in storage ($\pi^{[\mathrm{s}]}$)	£0.01/W
Circuit power consumption at the n-th BS ($P_n^{[\mathrm{c}]}$)	30 dBm
Maximum transmit power allowance ($P_n^{[\mathrm{Tmax}]}$)	46 dBm
Fronthaul capacity limit at the n-th BS ($B_n^{[\mathrm{limit}]}$)	35 bits/s/Hz
Storage capacity upper limit at the n-th BS ($E_n^{[\mathrm{capacity}]}$)	30 dBm
The discount factor in Step 13 of Algorithm 6.3.1 ($\mathcal{D}$)	0.95
Total number of time slots (T)	62
Total number of learning trials in each time slot (K)	7
In-advance energy charging packages offered at the grid	$\{100, 200, \cdots, 2000\}$ mW

Update adjusted reward $\bar{\mathbf{r}}_{\mathbf{n}}^{[\mathbf{t}]} = (\bar{r}_{n,1}^{[t]}, \bar{r}_{n,2}^{[t]}, \ldots, \bar{r}_{n,J}^{[t]})$, *as* $\bar{r}_{n,e}^{[t]} = \hat{r}_{n,e}^{[t]} + \sqrt{\frac{3lnt}{2N_e(t)}}$,

where $N_e(t)$ *is number of times the* e*-th arm has been played by the* t*-th time slot,*

24: **else if** t *is* **Exploitation**

25: *Solve problem in* (6.8),

26: **end if**

27: Average $\bar{\mathbf{r}}_{\mathbf{n}}^{[\mathbf{t}]}$ *over accumulated number of time slots, as*

$\bar{\mathbf{r}}_{\mathbf{n}} = \frac{\sum_{t'=1}^{t} \bar{\mathbf{r}}_{\mathbf{n}}^{[\mathbf{t}']} \mathcal{D}^{(t-t')}}{t} = [\bar{r}_{n,1}, \bar{r}_{n,2}, \cdots, \bar{r}_{n,J}], n \in \mathcal{L}_b.$

28: For the next time slot: find N *optimum arm indexes as*

$$e_n^* = \arg\max_e(\bar{r}_{n,e}), e \in \mathcal{J}, \forall n \in \mathcal{L}_b.$$

29: **End for**

6.4 Simulation Results

Similar to Section 5.5, this chapter considers a coordinated multipoint network of 3 adjacent 8-antenna BSs serving 6 single-antenna UTs and adopts a correlated channel model $\mathbf{h}_{ni} = \mathbf{C}_{ni}^{1/2}\mathbf{h}_w$ [1]. The renewable energy supply at BSs at each time slot varies as $G_1 \in [0.5\ 1.0]$ W, $G_2 \in [0.1\ 0.5]$ W and $G_3 \in [0.03\ 0.1]$ W, respectively, at $\pi^{[g]} = £0.05$/W. It is assumed in this chapter that one exploration time slot is followed by 3 exploitation time slots. The proposed algorithm is simulated with $K = 7$ trials averaging over $F = 20$ independent channel realizations for $T = 62$ time slots. The simulation parameters are summarized in Table 6.1 [4–7].

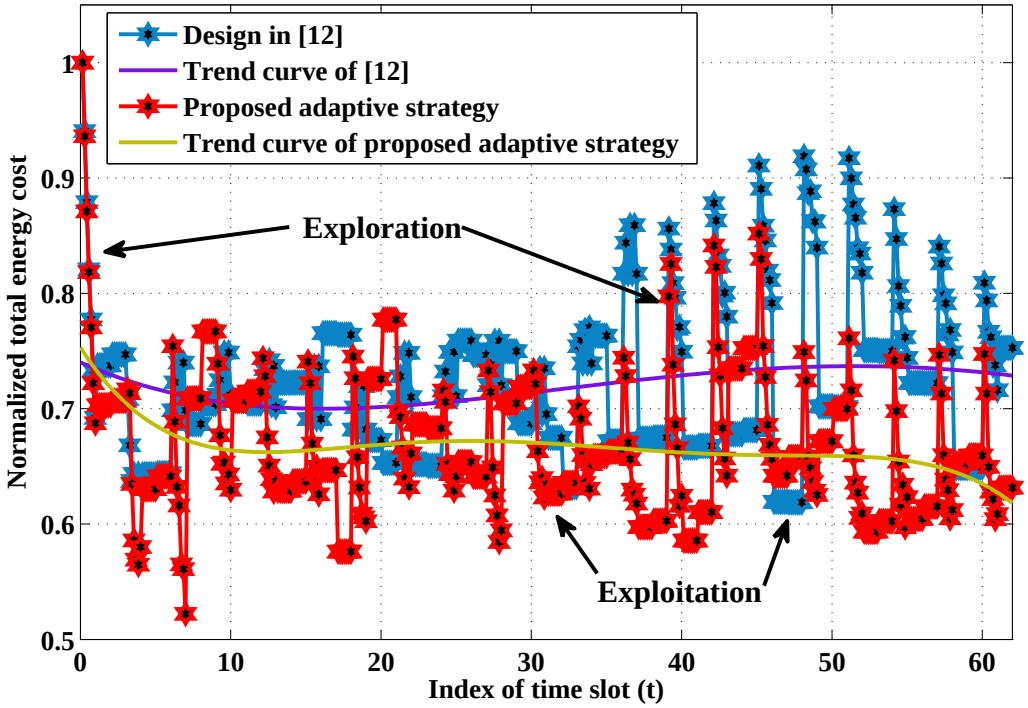


Figure 6.3: Normalized total energy cost of the proposed strategy versus design in [12] at $\gamma = 15$ dB at individual time slots

The normalized total energy cost of the proposed strategy at $\gamma = 15$ dB is compared in Fig. 6.3 against a simplified CMAB based storage-free energy

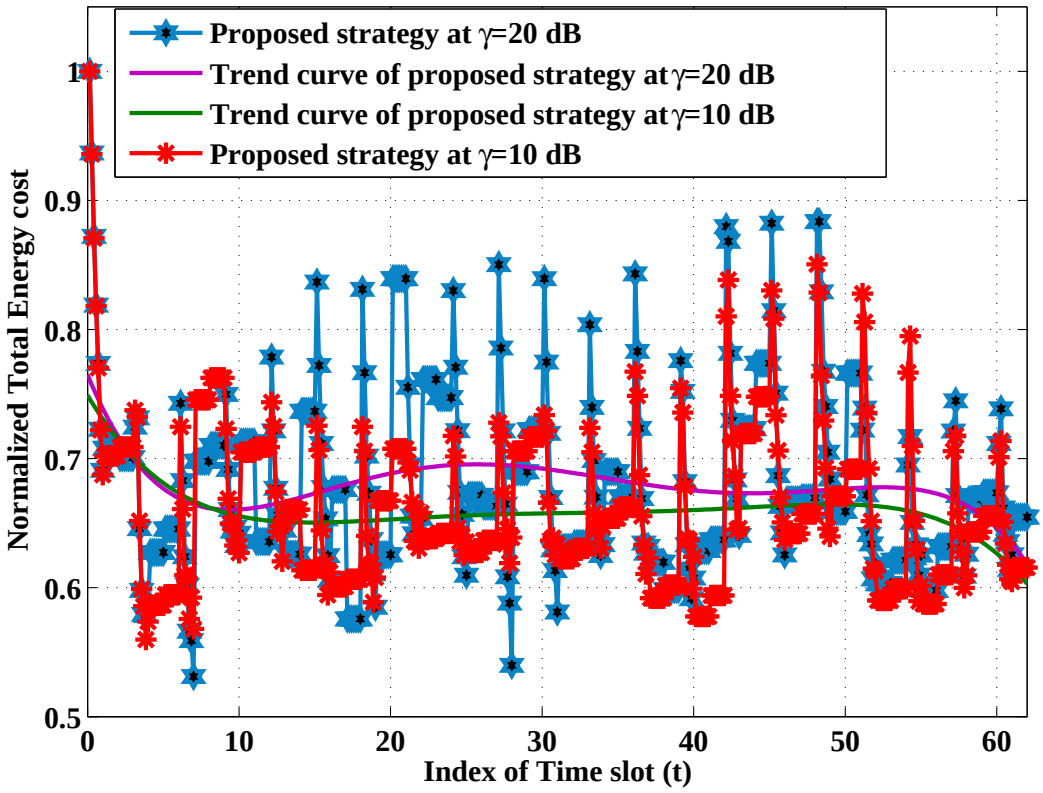


Figure 6.4: Normalized total energy cost of proposed strategy at $\gamma = 10$ dB and $\gamma = 20$ dB at individual time slots

management design in [12] that take no consideration of the randomness of the renewable energy generation, the wireless network dynamics, the long-term effect or the deployment of energy storage devices, and only explores the new super arms in an increasing order. For fair comparison, identical constraints are applied to all strategies and the overall energy cost is normalized to the energy cost at the initial trial of the proposed algorithm. The polynomial trend curves, fitted onto the actual experimental points, is adopted to better evaluate the average performance and the convergence of the proposed strategy. The burst at the start of an exploration cycle is due to the uncertain renewable energy generation and the perturbation in step 23 of Algorithm 6.3.1 to give priority to explore the less-explored arms. As shown in Fig. 6.3, the fitted polynomial trend curves approximately indicate that the averaged

performance of the proposed strategy achieves, respectively, 34 percent and 10 percent improvements over its initial learning state and the design in [12]. Furthermore, as the time-slot index increases, the design in [12] indicates larger variations in total energy cost and worse average performance than the proposed strategy. This is due to the single directional exploration and the storage-free nature of the design in [12], which provides poorer adaptation to the wireless channel dynamics and variations in renewable generation.

The proposed algorithm is evaluated in terms of normalized total energy cost at two more different SINR targets of $\gamma = 10$ dB and $\gamma = 20$ dB in Fig. 6.4. It is shown that the average performance of the proposed algorithm slightly degrades, i.e., has a larger total energy cost variation range and a polynomial trend curve with higher normalized total energy cost, as the target SINR increases within a substantial dynamic range, i.e., from $\gamma = 10$ dB to $\gamma = 20$ dB.

6.5 Concluding Remarks

The variability of renewable sources introduces large ramps in energy supply and significant fluctuations on the electricity price as well as grid stability issues. Addressing these issues, this chapter studies the problem of adaptive energy storage management in green wireless networks in the presence of uncertain renewable energy generation and dynamic wireless channel environment. A CMAB model is adopted to formulate the problem as a combination of online learning and optimal cost-aware energy coordination amongst the BSs to minimize the network cost over an infinite time horizon. A storage management algorithm is introduced to address the uncertain variations in energy supply and energy prices via adaptive power balancing at BSs. Simulation results confirm the effectiveness of the proposed learning-based storage management strategy in achieving an approximately 10 percent performance improvement over a recently proposed storage-free learning-based design in [12].

Chapter 7

Conclusions and future work

7.1 Thesis Summary

The ever-increasing energy consumption incurred by next generation dense wireless communication networks has always been considered as one of the most challenging issues from both ecological and economic perspectives. This thesis focuses on learning based energy management for green communications in multi-cell interference networks from two perspectives. From the first perspective, the cross-link coupling effect among a cluster of base stations (BSs), e.g., intercell interference (ICI), is taken into consideration and the alternatives to the existing coordinated transmission strategies and the robustness against the imperfect channel state information (CSI) are examined in Chapter 3 and 4. From the second perspective, dynamic nature of both wireless networks and renewable energy generation have been taken into account, and reinforcement learning based algorithms are proposed Chapter 5 and 6 to achieve a reliable and cost-efficient operation of the BSs supplied by a hybrid grid/renewable energy generators.

Chapter 1 outlines the motivation, contributions and the structure of this thesis. Chapter 2 provides a literature survey of the downlink energy management in multi-cell interference networks and the recent advances in robust beamforming, cooperative transmission and reinforcement learning. Furthermore, the mathematical preliminaries used in the subsequent chapters such as convex optimization are also introduced in this chapter.

In Chapter 3, two robust distributed coordinated transmission strategies that minimize the aggregate downlink transmit power in the presence of imperfect CSI in multi-cell interference networks are studied. Due to the fact that worst-case is a rare occurrence in practical network, the problems are constrained to satisfying a set of signal-to-interference-plus-noise-ratio (SINR) requirements and providing robustness against the second order statistical and instantaneous CSI uncertainties at individual user terminals (UTs) with certain SINR outage probabilities, respectively. The multicell-wise intractable optimization problems are first converted to the tractable form with linear matrix inequality constraints in a centralized manner, and then, decomposed into a set of independent parallel subproblems at individual BSs. The proposed iterative subgradient algorithm allows the individual BSs to iteratively learn transmit power level of each other and coordinate ICI among the BSs with a light inter-BS communication overhead. Simulation results demonstrate the advantages of these two proposed outage based probabilistic distributed transmission strategies in terms of providing larger SINR operational range as compared with worst-case robust beamforming designs in [10, 50] and outage probability based robust beamforming design in [9]. Besides, in terms of power efficiency, the proposed strategies have approximately 5% performance improvement as compared to the worst-case designs in [50] and [10] up to medium SINR operational range.

Chapter 4 introduces a distributed robust approach for maximizing the weighted SINR requirements at UTs in the presence of imperfect CSI in multi-cell interference networks, where the worst-case deterministic model is adopted for CSI imperfection. The optimization is constrained to strict individual BS transmit power limitations. Instead of solving the optimization problem directly, the original problem is converted into an equivalent total transmit power minimization problem based on the inverse relationship between the max-min SINR problem and the sum-power minimization problem. Taking account the cross-link coupling effect among BSs, an upper confidence bound based algorithm is proposed for the individual BSs to distributively learn the optimal achievable percentage coefficient of SINR targets

based on per BS power restrictions, and coordinate ICI across the BSs via light inter-BS communications. Simulation results confirm that the proposed strategy provides larger SINR operation range as compare to the centralized robust design in [51] and the distributed robust design in [50], as it always provides a feasible solution at the scaled target SINR.

In Chapter 5, a combinatorial multi-armed bandit (CMAB)-inspired online learning algorithm is introduced to account for the wireless channel random dynamism and minimize the time-averaged energy cost at individual BSs, powered by various energy markets and local renewable energy sources, over a finite time horizon. The proposed strategy benefits from an efficient trade-off between the exploration (i.e., online learning) and the exploitation (i.e., operational) modes, and sustains traffic demands by enabling sparse beamforming to schedule dynamic user-to-BS allocation and proactive energy provisioning at BSs to make ahead-of-time price-aware energy management decisions. Simulation results validate that in terms of reducing the overall energy cost, an average performance percentage improvement of 40, 8 and 7 per cent can be achieved by the proposed strategy as compared with recently proposed non-learning based designs in [4, 26] and a simplified CMAB based design in [12], respectively.

In Chapter 6, a CMAB-inspired online learning strategy is proposed for adaptive energy storage management and cost-aware coordinated load control at BSs to address the dynamic statistics of green wireless networks as well as the variability of renewable energy supply that are practically unknown in advance. The proposed strategy makes online foresighted decisions on the amount of energy to be stored in storage, such that the average energy cost over long time horizon can be minimized. It has been illustrated from the simulation results that in terms of total energy cost, the proposed learning-based storage management strategy achieves an average performance improvement of approximately 10 percent over a recently proposed storage-free learning-based design in [12].

7.2 Future Research Directions

The results attained in this thesis suggest several interesting future research directions that are highlighted as follows,

The decentralized transmission strategies studied in Chapter 3 and Chapter 4 adopt the iterative subgradient learning algorithm to coordinate ICI among the BSs. In order to solve the sum-power minimization problem in a distributed manner, the BSs are constrained to gradually learn the ICI and circulate key intercell coupling parameters in multiple iterations via inter-BS communications. Consequently, applying online learning to ICI coordination in a decentralized fashion, where the individual BSs can forecast the transmit power levels of other BSs and react based on its prediction, is deemed to be worthy for further investigation.

Chapter 5 and Chapter 6 study the foresighted cost-efficient energy management designs for green communications in a centralized coordinated cluster of small cells. However, the designs provide no robustness against the CSI estimation errors, which may lead to inefficiency in energy management in a practical scenario and may severely affect the system performance. Therefore, one possible future research direction is the robust energy management design in a decentralized scenario, where the individual BSs are equipped with energy storage devices and act as microgrid, such that the excessive energy can also be traded to the BSs that are in power shortage and the overall energy cost can be further reduced.

Furthermore, Chapter 5 and Chapter 6 focus merely on a single coordinated cluster and the impacts of individual clusters and/or the individual network operators on the global-wide electrical grid in the green wireless networks have been neglected. Thus, another future research direction could be the game theoretical approach to the global-wide cost-efficient energy management. More specifically, the penalty can be applied to the players with high energy consumption that influence more the electricity price. On the contrary, the incentive scheme can be involved to motivate the selfless and low energy consumed players, such that the cost-efficiency of the entire network can be achieved from a rather macroscopic perspective.

In addition, Chapter 5 proposes smart scheduling that reduces the fraction of exploration with increasing time slots. However, in the case of highly dynamic environment, the proposed smart scheduling may not be able to track and learn the fast changes in the environment. Thus, the adaptive ϵ-greedy method, e.g., the value-difference based exploration method [82], can be employed in the future research to adapt the exploration-exploitation trade-off to the uncertainty in the learning progress. More specifically, a time-decayed exploration rate can be adopted in a relative static environment, where the estimation of the mean reward process of the arms is improved with time and thus the high-cost exploration cycle can be reduced. On the contrary, a relative high exploration rate can be employed when a sudden change in the environment or the reward is observed.

Finally, the aforementioned studies are based on the assumption of single-antenna UTs. Future researches could be extended to multi-antennas UTs, e.g., massive multiple input multiple output (MIMO), which is deemed to be a promising solution for significant performance improvement in next generation dense networks. However, it necessitates both transmit and receive beamforming to be jointly designed, which arose several new challenges for the massive MIMO such as the need of efficient acquisition scheme for CSI as well as the significant increased the complexity and energy consumption of the signal processing at both the transmitters and the receivers. Thus, practical solutions to the optimal beamforming and trade-off between optimality and complexity are open problems for research.

Appendix A

Proof of Lemma 3.2.1

Following similar steps as in [58], a proof for Lemma 3.2.1 will be provided in the sequel. Let us start by rewriting $\mathrm{tr}(\mathbf{L}\boldsymbol{\Delta})$ in Lemma 3.2.1 as $\mathrm{tr}(\mathbf{L}\boldsymbol{\Delta}) = (\mathrm{vec}(\mathbf{L}^H))^H \mathrm{vec}(\boldsymbol{\Delta})$. Since $\mathrm{tr}(\mathbf{L}\boldsymbol{\Delta})$ can be recast as a linear combination of independently distributed zero-mean circularly symmetric complex Gaussian (ZMCSCG) random variables, $\mathrm{tr}(\mathbf{L}\boldsymbol{\Delta})$ is also a ZMCSCG random variable and can be characterized as $\mathrm{tr}(\mathbf{L}\boldsymbol{\Delta}) \sim \mathbb{CN}(0, \sigma^2_{\mathbf{L}\boldsymbol{\Delta}})$. $\sigma^2_{\mathbf{L}\boldsymbol{\Delta}}$ can be expressed as follows,

$$\begin{aligned}
\sigma^2_{\mathbf{L}\boldsymbol{\Delta}} &= \mathbb{E}[(\mathrm{vec}(\mathbf{L}^H))^H \mathrm{vec}(\boldsymbol{\Delta})\mathrm{vec}(\boldsymbol{\Delta})^H \mathrm{vec}(\mathbf{L}^H)] \\
&= (\mathrm{vec}(\mathbf{L}^H))^H \mathbb{E}[\mathrm{vec}(\boldsymbol{\Delta})\mathrm{vec}(\boldsymbol{\Delta})^H]\mathrm{vec}(\mathbf{L}^H) \\
&= (\mathrm{vec}(\mathbf{L}^H))^H \mathrm{diag}[\mathrm{vec}({\boldsymbol{\Sigma}_{\boldsymbol{\Delta}}}^H)]\mathrm{diag}[\mathrm{vec}(\boldsymbol{\Sigma}_{\boldsymbol{\Delta}})]\mathrm{vec}(\mathbf{L}^H) \\
&= (\mathrm{diag}[\mathrm{vec}(\boldsymbol{\Sigma}_{\boldsymbol{\Delta}})]\mathrm{vec}(\mathbf{L}^H))^H \mathrm{diag}[\mathrm{vec}(\boldsymbol{\Sigma}_{\boldsymbol{\Delta}})]\mathrm{vec}(\mathbf{L}^H) \\
&= \|\mathcal{D}_{\boldsymbol{\Delta}} \mathrm{vec}(\mathbf{L})\|^2,
\end{aligned}$$

where $\mathcal{D}_{\boldsymbol{\Delta}} = \mathrm{diag}(\mathrm{vec}({\boldsymbol{\Sigma}_{\boldsymbol{\Delta}}}^H))$. Hence proved $\mathrm{tr}(\mathbf{L}\boldsymbol{\Delta}) \sim \mathbb{N}(0, \|\mathcal{D}_{\boldsymbol{\Delta}} \mathrm{vec}(\mathbf{L})\|^2)$. Let $U \sim \mathbb{N}(0,1)$ be the standard normal random variable, then $\mathrm{tr}(\mathbf{L}\boldsymbol{\Delta}) \sim \mathbb{N}(0, \|\mathcal{D}_{\boldsymbol{\Delta}} \mathrm{vec}(\mathbf{L})\|^2)$ is equivalent to $\mathrm{tr}(\mathbf{L}\boldsymbol{\Delta}) = \|\mathcal{D}_{\boldsymbol{\Delta}} \mathrm{vec}(\mathbf{L})\| U$, $U \sim \mathbb{N}(0,1)$.

Appendix B

Proof of Lemma 4.3.1

In the sequel, a proof for Lemma 4.3.1 in the context of optimization problem in (4.14) will be provided on the basis of the Karush-Kuhn-Tucker (KKT) conditions. The Lagrangian of the optimization problem in (4.14) is given in (4.16) in Chapter 4. Noticing that $\mathbf{R}_{iik} = \frac{1}{r_e^2}\mathbf{I}_M$, let us start by rewriting $\mathbf{E}_{ik}$ and $\mathbf{F}_{ijk}$ in (4.14) as

$$\mathbf{E}_{ik} = \mathbf{\Lambda}_{ik} + \hat{\mathbf{H}}_{iik}^H \mathbf{\Phi}_{ik} \hat{\mathbf{H}}_{iik}, \tag{B.1}$$

$$\mathbf{F}_{ijk} = \mathbf{\Lambda}_{ijk} - \hat{\mathbf{H}}_{ijk}^H \mathbf{\Psi}_{ik} \hat{\mathbf{H}}_{ijk}, \tag{B.2}$$

where

$$\begin{cases} \mathbf{\Lambda}_{ik} &= \begin{bmatrix} \frac{\mu_{ik}}{r_e^2}\mathbf{I}_M & 0 \\ 0 & -\sigma_{ik}^2 - \mu_{ik} - \mathbf{d}_{iik}^T\mathbf{p} \end{bmatrix}, \\ \hat{\mathbf{H}}_{iik} &= \begin{bmatrix} \mathbf{I}_M & \hat{\mathbf{h}}_{iik} \end{bmatrix}, \\ \mathbf{\Lambda}_{ijk} &= \begin{bmatrix} \frac{\mu_{ijk}}{r_e^2}\mathbf{I}_M & 0 \\ 0 & -\mu_{ijk} - \mathbf{d}_{ijk}^T\mathbf{p} \end{bmatrix}, \\ \hat{\mathbf{H}}_{ijk} &= \begin{bmatrix} \mathbf{I}_M & \hat{\mathbf{h}}_{ijk} \end{bmatrix}, \end{cases}$$

and $\hat{\mathbf{H}}_{iik}, \hat{\mathbf{H}}_{ijk} \in \mathbb{C}^{M\times(M+1)}$. The KKT conditions are given by

$$\bigtriangledown_{\mathbf{w}_{ik}} L_i = 0, \tag{B.3}$$

$$\mathbf{E}_{ik}\lambda_{ik} = 0, \tag{B.4}$$

$$\mathbf{W}_{ik}\mathbf{A}_{ik} = 0, \tag{B.5}$$

$$\mathbf{A}_{ik} \succeq 0, \lambda_{ik} \succeq 0, \lambda_{ijk} \succeq 0, \ \forall k. \tag{B.6}$$

Then, by substituting (B.1) and (B.2) to (B.3), we can obtain

$$\mathbf{A}_{ik} = \mathbf{B}_{ik} - \hat{\mathbf{H}}_{iik}\lambda_{ik}(c_i\gamma_{ik})^{-1}\hat{\mathbf{H}}_{iik}^H, \tag{B.7}$$

where

$$\mathbf{B}_{ik} = \quad \mathbf{I}_M + \sum_{\substack{n\neq k,\\ n\in\mathcal{L}_i}} \hat{\mathbf{H}}_{iik}\lambda_{in}\hat{\mathbf{H}}_{iik}^H + \sum_{\substack{j\neq i,\\ j\in\mathcal{L}_b}} \sum_{k\in\mathcal{L}_i} \hat{\mathbf{H}}_{ijk}\lambda_{ijk}\hat{\mathbf{H}}_{ijk}^H. \tag{B.8}$$

In the sequel, it will be proved that rank $(\hat{\mathbf{H}}_{iik}\lambda_{ik}\hat{\mathbf{H}}_{iik}^H) \leq 1$. By substituting (B.1) into (B.4) and post-multiplying $\hat{\mathbf{H}}_{iik}^H$ on both sides of (B.4), we have the following expression

$$\mathbf{\Lambda}_{ik}\lambda_{ik}\hat{\mathbf{H}}_{iik}^H + \hat{\mathbf{H}}_{iik}^H\mathbf{\Phi}_{ik}\hat{\mathbf{H}}_{iik}\lambda_{ik}\hat{\mathbf{H}}_{iik}^H = 0. \tag{B.9}$$

Then, by pre-multiplying $[\mathbf{I}_M\ 0]$ on both sides of (B.9), we can obtain

$$\begin{aligned} & [\mathbf{I}_M\ 0]\mathbf{\Lambda}_{ik}\lambda_{ik}\hat{\mathbf{H}}_{iik}^H + [\mathbf{I}_M\ 0]\hat{\mathbf{H}}_{iik}^H\mathbf{\Phi}_{ik}\hat{\mathbf{H}}_{iik}\lambda_{ik}\hat{\mathbf{H}}_{iik}^H \qquad (B.10)\\ = \quad & \frac{\mu_{ik}}{r_e^2}[\mathbf{I}_M\ 0]\lambda_{ik}\hat{\mathbf{H}}_{iik}^H + \mathbf{I}_M\mathbf{\Phi}_{ik}\hat{\mathbf{H}}_{iik}\lambda_{ik}\hat{\mathbf{H}}_{iik}^H = 0\\ = \quad & \frac{\mu_{ik}}{r_e^2}(\hat{\mathbf{H}}_{iik}^H - [0_M\ \hat{\mathbf{h}}_{iik}])\lambda_{ik}\hat{\mathbf{H}}_{iik}^H + \mathbf{I}_M\mathbf{\Phi}_{ik}\hat{\mathbf{H}}_{iik}\lambda_{ik}\hat{\mathbf{H}}_{iik}^H. \end{aligned}$$

After simple mathematical deviation, it can be obtained that

$$\frac{\mu_{ik}}{r_e^2}[0_M\ \hat{\mathbf{h}}_{iik}]\lambda_{ik}\hat{\mathbf{H}}_{iik}^H = (\frac{\mu_{ik}}{r_e^2}\mathbf{I}_M + \mathbf{\Phi}_{ik})\hat{\mathbf{H}}_{iik}\lambda_{ik}\hat{\mathbf{H}}_{iik}^H. \tag{B.11}$$

By noticing the fact that the Hermitian matrix $\mathbf{E}_{ik} \succeq 0$, it is clear as per (4.14) that $\frac{\mu_{ik}}{r_e^2}\mathbf{I}_M + \mathbf{\Phi}_{ik} \succeq 0$ and it is nonsingular. Due to the fact that multiplying by a nonsingular matrix does not change the rank of a matrix, the following inequality

can be obtained as per rank properties.

$$\begin{aligned}
&\text{rank } (\hat{\mathbf{H}}_{iik}\lambda_{ik}\hat{\mathbf{H}}_{iik}^H) = \text{rank } ((\frac{\mu_{ik}}{r_e^2}\mathbf{I}_M + \mathbf{\Phi}_{ik})\hat{\mathbf{H}}_{iik}\lambda_{ik}\hat{\mathbf{H}}_{iik}^H) \\
&= \text{rank } (\frac{\mu_{ik}}{r_e^2}[0_M \ \hat{\mathbf{h}}_{iik}]\lambda_{ik}\hat{\mathbf{H}}_{iik}^H) \\
&\leq \min\Big(\text{rank } ([0_M \ \hat{\mathbf{h}}_{iik}]), \text{rank } (\lambda_{ik}\hat{\mathbf{H}}_{iik}^H)\Big) \\
&= \text{rank } ([0_M \ \hat{\mathbf{h}}_{iik}]) \leq 1.
\end{aligned} \tag{B.12}$$

In addition, according to the rank properties and (B.7), the following can be obtained

$$\begin{aligned}
\text{rank } (\mathbf{B}_{ik}) = \quad & \text{rank } (\mathbf{B}_{ik} + \hat{\mathbf{H}}_{iik}\lambda_{ik}(c_i\gamma_{ik})^{-1}\hat{\mathbf{H}}_{iik}^H \\
& -\hat{\mathbf{H}}_{iik}\lambda_{ik}(c_i\gamma_{ik})^{-1}\hat{\mathbf{H}}_{iik}^H) \\
\leq \quad & \text{rank } (\mathbf{B}_{ik} - \hat{\mathbf{H}}_{iik}\lambda_{ik}(c_i\gamma_{ik})^{-1}\hat{\mathbf{H}}_{iik}^H) \\
& +\text{rank } (\hat{\mathbf{H}}_{iik}\lambda_{ik}(c_i\gamma_{ik})^{-1}\hat{\mathbf{H}}_{iik}^H) \\
= \quad & \text{rank } (\mathbf{A}_{ik}) + \text{rank } (\hat{\mathbf{H}}_{iik}\lambda_{ik}(c_i\gamma_{ik})^{-1}\hat{\mathbf{H}}_{iik}^H).
\end{aligned} \tag{B.13}$$

Thus, it can be concluded that

$$\begin{aligned}
\text{rank } (\mathbf{A}_{ik}) \geq \quad & \text{rank } (\mathbf{B}_{ik}) - \text{rank } (\hat{\mathbf{H}}_{iik}^H\lambda_{ik}(c_i\gamma_{ik})^{-1}\hat{\mathbf{H}}_{iik}) \\
\geq \quad & \text{rank } (\mathbf{B}_{ik}) - 1.
\end{aligned} \tag{B.14}$$

In the sequel, it will be shown by contradiction that $\mathbf{B}_{ik} \succ 0$ always holds. Assuming that $\mathbf{B}_{ik} \succeq 0$, there must exist a vector $\mathbf{a} \neq 0$ such that $\mathbf{a}^H\mathbf{B}_{ik}\mathbf{a} = 0$. Then the (B.7) can be rewritten as

$$\begin{aligned}
\mathbf{a}^H\mathbf{A}_{ik}\mathbf{a} \quad & = -\mathbf{a}^H(\hat{\mathbf{H}}_{iik}\lambda_{ik}(c_i\gamma_{ik})^{-1}\hat{\mathbf{H}}_{iik}^H)\mathbf{a} \\
& = -(c_i\gamma_{ik})^{-1}|\mathbf{a}^H\hat{\mathbf{H}}_{iik}\lambda_{ik}^{\frac{1}{2}}|^2 < 0,
\end{aligned} \tag{B.15}$$

which indicates that $\mathbf{A}_{ik}$ is not positive-definite and contradicts to (B.6). Hence, $\mathbf{B}_{ik} \succ 0$ always holds and rank $(\mathbf{A}_{ik}) = M$ or rank $(\mathbf{A}_{ik}) = M - 1$, provided that

rank $(\mathbf{B}_{ik}) = M$. Furthermore, in accordance with the KKT condition in (B.5), the columns of $\mathbf{W}_{ik}$ are in the null space of $\mathbf{A}_{ik}$, i.e., rank $(\mathbf{W}_{ik}) = 1$ holds if rank $(\mathbf{A}_{ik}) = M-1$. However, if rank $(\mathbf{A}_{ik}) = M$, then $\mathbf{W}_{ik} = 0$, which indicates that $\mathbf{W}_{ik}$ is not an optimal solution to the problem in (4.14). Thus, rank $(\mathbf{A}_{ik}) = M - 1$ and it can be easily concluded that rank $(\mathbf{W}_{ik}) = 1$. Hence, rank $(\mathbf{W}_{ik}) = 1$ holds with probability one. This thus completes the proof of Lemma 4.3.1 for problem (4.14).

Appendix C

Proof of Lemma 5.3.1

Following similar steps as in [4], a proof for Lemma 5.3.1 in the context of optimization problem in (5.9) will be provided in the sequel. For the sake of notational simplicity, let us denote the aggregate beamforming and channel vectors from all the BSs towards the i-th UT, $i \in \mathcal{L}_i$, as $\mathbf{w}_i = [\mathbf{w}_{1i}^H, \cdots, \mathbf{w}_{Ni}^H]^H \in \mathbb{C}^{MN\times 1}$ and $\mathbf{h}_i = [\mathbf{h}_{1i}^H, \cdots, \mathbf{h}_{Ni}^H]^H \in \mathbb{C}^{MN\times 1}$, respectively. Let us further define a block diagonal matrix $\mathbf{D}_n \triangleq \text{Bdiag}(\mathbf{0}_1 \cdots \mathbf{0}_i \ldots \mathbf{I}_n \cdots \mathbf{0}_N) \succeq 0, \forall n \in \mathcal{L}_b$, such that $\text{tr}(\mathbf{W}_{ni}) = \text{tr}(\mathbf{W}_i\mathbf{D}_n)$, where $\mathbf{W}_i = \mathbf{w}_i\mathbf{w}_i^H$ is a rank-one semidefinite matrix. Then, the convex optimization problem in (5.9) can be recast as follows,

$$
\begin{aligned}
\min_{\mathbf{W}_i, E_n^{[\mathrm{r}]}(t), S_n(t)} \quad & \sum_{i\in\mathcal{L}_i} \text{tr}(\mathbf{W}_i) + \sum_{n\in\mathcal{L}_b} \left\{E_n^{[\mathrm{r}]}(t)\right\} \qquad \text{(C.1)}\\
\text{s.t.} \quad & \text{C1}: \gamma_i^{-1}\text{tr}(\mathbf{H}_i\mathbf{W}_i) \geq \sum_{j\in\mathcal{L}_i, j\neq i} \text{tr}(\mathbf{H}_i\mathbf{W}_j) + \sigma_i^2, \forall i \in \mathcal{L}_i,\\
& \text{C2}: \sum_{i\in\mathcal{L}_i} \xi_{ni}\text{tr}(\mathbf{W}_i\mathbf{D}_n)R_i \leq B_n^{[\mathrm{limit}]},\ \forall n \in \mathcal{L}_b,\\
& \text{C3}: \eta\sum_{i\in\mathcal{L}_i} \text{tr}(\mathbf{W}_i\mathbf{D}_n) \leq G_n(t) + E_n^{[\mathrm{a}]}(t) - P_n^{[\mathrm{c}]} - S_n(t) + E_n^{[\mathrm{r}]}(t),\\
& \text{C4}: \sum_{i\in\mathcal{L}_i} \text{tr}(\mathbf{W}_i\mathbf{D}_n) \leq P_n^{[\mathrm{Tmax}]}, \forall n \in \mathcal{L}_b,\\
& \text{C5}: E_n^{[\mathrm{r}]}(t) \geq 0, \quad \forall n \in \mathcal{L}_b,\\
& \text{C6}: S_n(t) \geq 0, \quad \forall n \in \mathcal{L}_b,\\
& \text{C7}: \mathbf{W}_i \succeq 0, \quad \forall i \in \mathcal{L}_i.
\end{aligned}
$$

In the sequel, it will be shown by contradiction that $\text{rank}(\mathbf{W}_i^*) \leq 1$ holds with probability one. For simplicity, the index t is omitted for the rest of the proof. The

convex optimization problem in (C.1) satisfies the Slater's condition, thus, strong duality holds [8]. Let us define $\mathbf{Y}_i$ and the set $\Theta = \{\nu_i, \varphi_n, \phi_n, \tau_n, \epsilon_n, \varrho_n\}$, respectively, as the dual variable matrix of C7 and the set of scalar Lagrange multipliers of constraints C1-C6. The Lagrangian of the optimization problem in (C.1) can then be expressed as

$$\begin{aligned} & \mathcal{L}(\mathbf{W}_i, E_n^{[\mathrm{r}]}, S_n, \mathbf{Y}_i, \nu_i, \varphi_n, \phi_n, \tau_n, \epsilon_n, \varrho_n) \\ = \quad & \sum_{i\in\mathcal{L}_i} \mathrm{tr}(\mathbf{Q}_i\mathbf{W}_i) - \sum_{i\in\mathcal{L}_i} \mathrm{tr}(\mathbf{W}_i(\mathbf{Y}_i + \frac{\nu_i\mathbf{H}_i}{\gamma_i})) + \Xi, \end{aligned} \tag{C.2}$$

where

$$\mathbf{Q}_i = \mathbf{I} + \sum_{j\in\mathcal{L}_i, j\neq i} \nu_j\mathbf{H}_j + \sum_{n\in\mathcal{L}_b} (\eta\phi_n + \tau_n + \varphi_n\xi_{ni}R_i)\mathbf{D}_n, \tag{C.3}$$

and

$$\begin{aligned} \Xi \quad = & \sum_{n\in\mathcal{L}_b} E_n^{[\mathrm{r}]} - \sum_{n\in\mathcal{L}_b} (\phi_n + \epsilon_n)E_n^{[\mathrm{r}]} + \sum_{n\in\mathcal{L}_b} (\phi_n - \varrho_n)S_n + \sum_{i\in\mathcal{L}_i} \nu_i\sigma_i^2 \\ & - \sum_{n\in\mathcal{L}_b} \varphi_n B_n^{[\mathrm{limit}]} - \sum_{n\in\mathcal{L}_b} \tau_n P_n^{[\mathrm{Tmax}]} - \sum_{n\in\mathcal{L}_b} \phi_n(G_n + E_n^{[\mathrm{a}]} + P_n^{[\mathrm{c}]}), \end{aligned} \tag{C.4}$$

is the summation of terms of variables that are independent of $\mathbf{W}_i$. The dual problem of problem in (C.1) is then given by

$$\max_{\Theta\geq 0, \mathbf{Y}_i\succeq 0} \min_{\mathbf{W}_i, E_n^{[\mathrm{r}]}, S_n} \quad \mathcal{L}(\mathbf{W}_i, E_n^{[\mathrm{r}]}, S_n, \mathbf{Y}_i, \Theta), \tag{C.5}$$

where $\Theta \geq 0$ indicates that all of the scalar dual variables within the set Θ are non-negative. Let us define $\{\mathbf{W}_i^*, E_n^{[\mathrm{r}]*}, S_n^*\}$ and $\{\mathbf{Y}_i^*, \Theta^*\}$ as the sets of optimal primal and dual variables of (C.1), respectively. The dual problem in (C.5) can be written as

$$\min_{\mathbf{W}_i} \ \mathcal{L}(\mathbf{W}_i, E_n^{[\mathrm{r}]*}, S_n^*, \mathbf{Y}_i^*, \Theta^*), \tag{C.6}$$

and the Karush-Kuhn-Tucker (KKT) conditions are given by

$$\Theta^* \geq 0,\ \mathbf{Y}_i^* \succeq 0,\ \mathbf{Y}_i^*\mathbf{W}_i^* = 0,\ \forall i \in \mathcal{L}_i, \tag{C.7}$$

$$\mathbf{Q}_i^* - (\mathbf{Y}_i^* + \frac{\nu_i^*\mathbf{H}_i}{\gamma_i}) = 0,\ \forall i \in \mathcal{L}_i, \tag{C.8}$$

where $\mathbf{Q}_i^* = \mathbf{I} + \sum_{j\in\mathcal{L}_i, j\neq i} \nu_j^*\mathbf{H}_j + \sum_{n\in\mathcal{L}_b} (\eta\phi_n^* + \tau_n^* + \varphi_n^*\xi_{ni}R_i)\mathbf{D}_n$. Let us first prove by contradiction that $\mathbf{Q}_i^*$ is a positive definite matrix with probability one. Suppose $\mathbf{Q}_i^*$ is a non-positive definite matrix, then one of the optimal solutions of (C.6) can be chosen as $\mathbf{W}_i = \hbar\mathbf{w}_i\mathbf{w}_i^H$, where $\hbar > 0$ is a scaling parameter and $\mathbf{w}_i$ is the eigenvector corresponding to one of the non-positive eigenvalues of $\mathbf{Q}_i^*$. Substituting $\mathbf{W}_i = \hbar\mathbf{w}_i\mathbf{w}_i^H$ into (C.6) leads to

$$\begin{aligned}&\min_{\mathbf{W}_i}\ \mathcal{L}(\mathbf{W}_i, E_n^{[\mathrm{r}]*}, S_n^*, \mathbf{Y}_i^*, \Theta^*) \\ &= \sum_{i\in\mathcal{L}_i} \mathrm{tr}(\hbar\mathbf{Q}_i^*\mathbf{w}_i\mathbf{w}_i^H) - \hbar\sum_{i\in\mathcal{L}_i} \mathrm{tr}(\mathbf{w}_i^H(\mathbf{Y}_i^* + \frac{\nu_i^*\mathbf{H}_i}{\gamma_i})\mathbf{w}_i) + \Theta^*\end{aligned} \tag{C.9}$$

where $\sum_{i\in\mathcal{L}_i} \mathrm{tr}(\hbar\mathbf{Q}_i^*\mathbf{w}_i\mathbf{w}_i^H)$ is non-positive and $-\hbar\sum_{i\in\mathcal{L}_i} \mathrm{tr}(\mathbf{w}_i^H(\mathbf{Y}_i^* + \frac{\nu_i^*\mathbf{H}_i}{\gamma_i})\mathbf{w}_i) \to \infty$ if $\hbar \to \infty$, which results in the dual optimal value unbounded from below. However, the optimal value of the primal problem is non-negative, thus strong duality does not hold which induces a contradiction. Hence, $\mathbf{Q}_i^*$ is a positive definite matrix with probability one and $\mathrm{rank}(\mathbf{Q}_i^*) = MN$. According to (C.8) and properties of rank of matrix, the following inequality holds

$$\begin{aligned}&\mathrm{rank}(\mathbf{Q}_i^*) = MN = \mathrm{rank}(\mathbf{Y}_i^* + \frac{\nu_i^*\mathbf{H}_i}{\gamma_i}) \leq \mathrm{rank}(\mathbf{Y}_i^*) + \mathrm{rank}(\frac{\nu_i^*\mathbf{H}_i}{\gamma_i}) \\ &\Rightarrow \mathrm{rank}(\mathbf{Y}_i^*) \geq MN - \mathrm{rank}(\frac{\nu_i^*\mathbf{H}_i}{\gamma_i}).\end{aligned} \tag{C.10}$$

Thus, $\mathrm{rank}(\mathbf{Y}_i^*) = MN - 1$ or $\mathrm{rank}(\mathbf{Y}_i^*) = MN$. Furthermore, the KKT condition in (C.7), i.e., $\mathbf{Y}_i^*\mathbf{W}_i^* = 0$, indicates that for $\mathbf{W}_i^* \neq 0$, the columns of $\mathbf{W}_i^*$ are in the null space of $\mathbf{Y}_i^*$, and $\mathbf{W}_i^* \neq 0$ is required to satisfy the minimum SINR requirements in constraint C1 for $\gamma_i > 0$. Hence, $\mathrm{rank}(\mathbf{W}_i^*) = 1$ holds with probability one.

References

[1] T. A. Le and M. R. Nakhai, "An iterative algorithm for downlink multi-cell beamforming," *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Dec. 2011.

[2] 3GPP, "Tr 36.814 v9.0.0: Further advancements for e-utra physical layer aspects (release 9)," Mar. 2010. [Online]. Available: www.3gpp.org/dynareport/36814.htm

[3] A. Shaverdian and M. R. Nakhai, "Robust distributed beamforming with interference coordination in downlink cellular networks," *IEEE Transactions on Communications*, vol. 62, no. 7, pp. 2411–2421, Jun. 2014.

[4] W. N. S. F. Wan-Ariffin, X. Zhang, and M. R. Nakhai, "Sparse beamforming for real-time resource management and energy trading in green c-ran," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 2022–2031, Jul. 2017.

[5] D. W. K. Ng and R. Schober, "Resource allocation for coordinated multipoint networks with wireless information and power transfer," *IEEE Global Communications Conference (GLOBECOM)*, pp. 4281–4287, Dec. 2014.

[6] B. Dai and W. Yu, "Sparse beamforming and user-centric clustering for downlink cloud radio access network," *IEEE Access*, vol. 2, pp. 1326–1339, Oct. 2014.

[7] Y. Zhang and M. v. d. Schaar, "Structure-aware stochastic storage management in smart grids," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1098–1110, Dec. 2014.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[9] C. Shen, T. H. Chang, K. Wang, Z. Qiu, and C. Chi, "Chance-constrained robust beamforming for multi-cell coordinated downlink," *IEEE Global Communications Conference (GLOBECOM)*, pp. 4957–4962, Dec. 2012.

[10] C. Shen, T. H. Chang, K. Y. Wang, Z. Qiu, and C. Y. Chi, "Distributed robust multicell coordinated beamforming with imperfect csi: An admm approach,"

IEEE Transactions on Signal Processing, vol. 60, no. 6, pp. 2988–3003, Feb. 2012.

[11] Z. Xiang, M. Tao, and X. Wang, "Coordinated multicast beamforming in multicell networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 12–21, Jan. 2013.

[12] W. N. S. F. Wan-Ariffin, X. Zhang, and M. R. Nakhai, "Combinatorial multi-armed bandit algorithms for real-time energy trading in green C-RAN," *IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2016.

[13] P. Gandotra, R. K. Jha, and S. Jain, "Green communication in next generation cellular networks: A survey," *IEEE Access*, vol. 5, pp. 11 727–11 758, Jun. 2017.

[14] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. S. Thompson, I. Ku, C. X. Wang, T. A. Le, M. R. Nakhai, J. Zhang, and L. Hanzo, "Green radio: radio techniques to enable energy-efficient wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 46–54, Jun. 2011.

[15] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 55–62, Aug. 2011.

[16] GSMA, "Green power for mobile. the global telecom tower esco market," Dec. 2014. [Online]. Available: https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2015/01/140617-GSMA-report-draft-vF-KR-v7.pdf

[17] T. C. Group, "Smart 2020: Enabling the low carbon economy in the information age," Jun. 2008. [Online]. Available: https://www.theclimategroup.org/sites/default/files/archive/files/Smart2020Report.pdf

[18] C. X. Wang, F. Haider, X. Gao, X. H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5g wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[19] S. Buzzi, C.-L. I, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5g networks and challenges ahead," *IEEE Journal of Selected Areas in Communications*, vol. 34, no. 4, pp. 697–709, 2016.

[20] T. A. Le, S. Nasseri, A. Zarrebini-Esfahani, and M. R. Nakhai, "Power-efficient downlink transmission in multicell networks with limited wireless backhaul," *IEEE Wireless Communications*, vol. 18, no. 5, pp. 82–88, Oct. 2011.

[21] D. H. N. Nguyen and T. Le-Ngoc, "Multiuser downlink beamforming in multicell wireless systems: A game theoretical approach," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 3326–3338, Jul. 2011.

[22] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 13, no. 4, pp. 524–540, Nov. 2011.

[23] O. Ellabban, H. Abu-Rub, and F. Blaabjerg, "Renewable energy resources: Current status, future prospects and their enabling technology," *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 748–764, Aug. 2014.

[24] T. Han and N. Ansari, "Powering mobile networks with green energy," *IEEE Wireless Communications*, vol. 21, no. 1, pp. 90–96, Feb. 2014.

[25] R. E. H. Sims, H.-H. Rogner, and K. Gregory, "Carbon emission and mitigation cost comparisons between fossil fuel, nuclear and renewable energy resources for electricity generation," *Energy policy*, vol. 31, no. 13, pp. 1315–1326, Oct. 2003.

[26] J. Xu and R. Zhang, "Cooperative energy trading in CoMP systems powered by smart grids," *IEEE Global Communications Conference (GLOBECOM)*, pp. 2697–2702, Dec. 2014.

[27] X. Wang, Y. Zhang, T. Chen, and G. B. Giannakis, "Dynamic energy management for smart-grid-powered coordinated multipoint systems," *IEEE Journal of Selected Areas in Communications*, vol. 34, no. 5, pp. 1348–1359, May 2016.

[28] D. Niyato, X. Lu, and P. Wang, "Adaptive power management for wireless base stations in a smart grid environment," *IEEE Wireless Communications*, vol. 19, no. 6, pp. 44–51, Dec. 2012.

[29] X. Zhang, M. R. Nakhai, and W. N. S. F. Wan-Ariffin, "Robust chance-constrained distributed beamforming for multicell interference networks," *IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2016.

[30] X. Zhang and M. R. Nakhai, "A distributed algorithm for robust transmission in multicell networks with probabilistic constraints," *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Dec. 2016.

[31] X. Zhang, M. R. Nakhai, and W. N. S. F. Wan-Ariffin, "A multi-armed bandit approach to distributed robust beamforming in multicell networks," *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Dec. 2016.

[32] X. Zhang, M. R. Nakhai, and W. N. S. F. Wan-Ariffin, "A bandit approach to price-aware energy management in cellular networks," *IEEE Communications Letter*, vol. 21, no. 7, pp. 1609–1612, Jul. 2017.

[33] X. Zhang, M. R. Nakhai, and W. N. S. F. Wan-Ariffin, "Adaptive energy storage management in green wireless networks," *IEEE Signal Processing Letter*, vol. 24, no. 7, pp. 1044–1048, Jul. 2017.

[34] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1426–1438, Jul. 2006.

[35] M. Grant and S. Boyd, "Cvx: Matlab software for disciplined convex programming, version 2.1," Jun. 2015. [Online]. Available: http://cvxr.com/cvx/doc/CVX.pdf

[36] J. F. Sturm, "Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 625–653, 1999.

[37] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1996.

[38] Z. Q. Luo, W. K. Ma, A. M. C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, May 2010.

[39] A. B. Gershman, N. D. Sidiropoulos, Shahhazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 62–75, May 2010.

[40] E. Karipidis, N. Sidiropoulos, and Z.-Q. Luo, "Quality of service and max-min transmit beamforming to multiple cochannel multicast groups," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1268–1279, Mar. 2008.

[41] E. Song, Q. Shi, M. Sanjabi, R. Sun, and Z. Q. Luo, "Robust sinr-constrained miso downlink beamforming: When is semidefinite programming relaxation tight?" *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 3096–C2099, May 2011.

[42] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 126–135, Dec. 2014.

[43] I. Hwang, B. Song, and S. S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 20–27, Jun. 2013.

[44] S. Sun, Q. Gao, Y. Peng, Y. Wang, and L. Song, "Interference management through comp in 3gpp lte-advanced networks," *IEEE Wireless Communications*, vol. 20, no. 1, pp. 59–66, Feb. 2013.

[45] F. Gross, *Smart Antennas for Wireless Communications*. McGraw-Hill, 2005.

[46] W. N. S. F. Wan-Ariffin, X. Zhang, and M. R. Nakhai, "Sparse beamforming for real-time energy trading in CoMP-SWIPT networks," *IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2016.

[47] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in multi-cell ofdma systems with limited backhaul capacity," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3618–3631, Apr. 2012.

[48] X. He and Y. Wu, "Tight probabilistic sinr constrained beamforming under channel uncertainties," *IEEE Transactions on Signal Processing*, vol. 63, no. 13, pp. 3490–3505, Apr. 2015.

[49] A. Tajer, N. Prasad, and X. Wang, "Robust linear precoder design for multi-cell downlink transmission," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 235–251, Jan. 2011.

[50] H. Pennanen, A. Tolli, and M. Latva-aho, "Decentralized robust beamforming for coordinated multi-cell miso networks," *IEEE Signal Processing Letters*, vol. 21, no. 3, pp. 334–338, Mar. 2014.

[51] C. Shen, K. Y. Wang, T. H. Chang, Z. Qiu, and C. Y. Chi, "Worst-case sinr constrained robust coordinated beamforming for multicell wireless systems," *IEEE International Conference on Communications (ICC)*, pp. 1–5, Jun. 2011.

[52] Y. W. Huang, D. P. Palomar, and S. Z. Zhang, "Lorentz-positive maps and quadratic matrix inequalities with applications to robust miso transmit beamforming," *IEEE Transaction on Signal Processing*, vol. 61, no. 5, pp. 1121–1130, Mar. 2013.

[53] S. Nasseri and M. R. Nakhai, "Min-max robust transmit beamforming for power efficient quality of service guarantee," *IEEE Global Communications Conference (GLOBECOM)*, pp. 3360–3365, Dec. 2014.

[54] N. Vucic and H. Boche, "Robust qos-constrained optimization of downlink multiuser miso systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 714–725, Oct. 2008.

[55] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in mimo-relay networks with imperfect csi," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5473C–5484, Nov. 2011.

[56] M. Tshangini and M. R. Nakhai, "Second-order cone programming for robust downlink beamforming with imperfect csi," *IEEE Global Communications Conference (GLOBECOM)*, pp. 3474–3479, Dec. 2013.

[57] M. Tshangini and M. R. Nakhai, "Robust downlink beamforming with imperfect csi," *IEEE International Conference on Communications (ICC)*, pp. 4916–4920, Dec. 2013.

[58] S. Nasseri and M. R. Nakhai, "Robust interference management via outage-constrained downlink beamforming in multicell networks," *IEEE Global Communications Conference (GLOBECOM)*, pp. 3470–3475, Dec. 2013.

[59] S. Nasseri, M. R. Nakhai, and T. A. Le, "Chance constrained robust downlink beamforming in multicell networks," *IEEE Transactions on Mobile Computing*, vol. 15, no. 11, pp. 2682–2691, Jan. 2016.

[60] K. Y. Wang, A. M. C. So, T. H. Chang, W. K. Ma, and C. Y. Chi, "Outage constrained robust transmit optimization for multiuser miso downlinks: Tractable approximations by conic optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5690–5705, Sep. 2014.

[61] K. Y. Wang, T. H. Chang, W. K. Ma, A. M. C. So, and C. Y. Chi, "Probabilistic sinr constrained robust transmit beamforming: A bernstein-type inequality based conservative approach," *IEEE International Conference on Acoustic, Speech and Signal processing (ICASSP)*, pp. 3080–3083, May 2011.

[62] P. J. Chung, H. Du, and J. Gondzio, "A probabilistic constraint approach for robust beamforming with imperfect channel information," *IEEE Transaction on Signal Processing*, vol. 59, no. 6, pp. 2773C–2782, Mar. 2011.

[63] B. K. Chalise, S. Shahbazpanahi, A. Czylwik, and A. B. Gersham, "Robust downlink beamforming based on outage probability specifications," *IEEE Transaction on Wireless Communications*, vol. 6, no. 10, pp. 3498C–3503, Oct. 2007.

[64] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H. P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance and field trial results," *IEEE Communications Magazine*, vol. 49, no. 2, pp. 102–111, Feb. 2011.

[65] J. Lee, Y. Kim, H. Lee, B. L. Ng, D. Mazzarese, J. Liu, W. Xiao, and Y. Zhou, "Coordinated multipoint transmission and reception in lte-advanced systems," *IEEE Communication Magzine*, vol. 50, no. 11, pp. 44–50, Nov. 2012.

[66] M. Hong, R. Sun, H. Baligh, and Z. Q. Luo, "Joint base station clustering and beamformer design for partial coordinated transmission in heterogenous networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 226–240, Feb. 2013.

[67] J. Zhao, T. Q. S. Quek, and Z. Lei, "Coordinated multipoint transmission with limited backhaul data transfer," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2762–2775, Jun. 2013.

[68] Z. Zhao, M. Peng, Z. Ding, C. Wang, and H. V. Poor, "Cluster formation in cloud-radio access networks: Performance analysis and algorithms design," *IEEE International Conference on Communications (ICC)*, pp. 3903–3908, May 2015.

[69] P. Rost, C. J. Bernardos, A. D. Domenico, M. D. Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wubben, "Cloud technologies for flexible 5g radio access networks," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 68–76, May 2014.

[70] K. Chen, "C-ran: The road towards green ran," Oct. 2011. [Online]. Available: http://labs.chinamobile.com/cran/wp-content/uploads/CRAN-white-paper-v2-5-EN.pdf

[71] M. Peng, C. Wang, V. Lau, and H. V. Poor, "Fronthaul-constrained cloud radio access networks: insights and challenges," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 152–160, Apr. 2015.

[72] S. Bu, F. R. Yu, Y. Cai, and X. P. Liu, "When the smart grid meets energy-efficient communications: Green wireless cellular networks powered by the smart grid," *IEEE Transactions on Wireless Communications*, vol. 11, no. 8, pp. 3014–3024, Aug. 2012.

[73] R. G. Pratt, P. J. Balducci, M. C. W. Kintner-Meyer, T. F. Sanquist, C. Gerkensmeyer, K. P. Schneider, S. Katipamula, and T. J. Secrest, "The smart grid: an estimation of the energy and co2 benefits," Jan. 2010. [Online]. Available: https://energyenvironment.pnnl.gov/news/pdf/PNNL-19112-Revision-1-Final.pdf

[74] I. S. Bayram, M. Z. Shakir, M. Abdallah, and K. Qaraqe, "A survey on energy trading in smart grid," *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 258–262, Dec. 2014.

[75] S. Chen, N. B. Shroff, and P. Sinha, "Energy trading in the smart grid: from end-users perspective," *Asilomar Conference on Signals, Systems and Computers*, pp. 327–331, Nov. 2013.

[76] J. Leithon, T. J. Lim, and S. Sun, "Energy exchange among base stations in a cellular network through the smart grid," *IEEE International Conference on Communications (ICC)*, pp. 4036–4041, May 2014.

[77] J. Leithon, T. J. Lim, and S. Sun, "Online energy management strategies for base stations powered by the smart grid," *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 199–204, Oct. 2013.

[78] Y. Wang, W. Saad, Z. Han, H. V. Poor, and T. Basar, "A game-theoretic approach to energy trading in the smart grid," *IEEE Transactions on Smart Grid*, vol. 5, no. 3, pp. 1439–1450, May 2014.

[79] X. Wang, Y. Zhang, G. B. Giannakis, and S. Hu, "Robust smart-grid-powered cooperative multipoint systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6188–6199, Jun. 2015.

[80] R. S. Sutton and A. G. Barto, *Reinforcement Learning An Introduction*. The MIT Press, 1998.

[81] L. Gavrilovska, V. Atanasovski, I. Macaluso, and L. A. DaSilva, "Learning and reasoning in cognitive radio networks," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 4, pp. 1761–1777, 2013.

[82] M. Tokic, *KI 2010: Advances in Artificial Intelligence*. Springer, 2010.

[83] K. Liu and Q. Zhao, "Distributed learning in multi-armed bandit with multiple players," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5667–5681, Nov. 2010.

[84] W. Chen, Y. Wang, and Y. Yuan, "Combinatorial multi-armed bandit: General framework, results and applications," *International Conference on Machine Learning*, Jun. 2013.

[85] P. Blasco and D. Gunduz, "Learning-based optimization of cache content in a small cell base station," *IEEE International Conference on Communications (ICC)*, pp. 1–6, Jun. 2014.

[86] S. Maghsudi and E. Hossain, "Multi-armed bandits with application to 5g small cells," *IEEE Wireless Communications*, vol. 23, no. 3, pp. 64–73, Jun. 2016.

[87] S. Maghsudi and S. Stanczak, "Joint channel selection and power control in infrastructureless wireless networks: A multiplayer multiarmed bandit framework," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, pp. 4565–4578, Oct. 2015.

[88] S. Maghsudi and E. Hossain, "Distributed user association in energy harvesting dense small cell networks: A mean-field multi-armed bandit approach," *IEEE Access*, vol. 5, pp. 3513–3523, Mar. 2017.

[89] S. Maghsudi and S. Stanczak, "Channel selection for network-assisted D2D communication via no-regret bandit learning with calibrated forecasting," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1309–1322, Mar. 2015.

[90] R. Estrada, A. Jarray, H. Otrok, Z. Dziong, and H. Barada, "Energy-efficient resource-allocation model for OFDMA macrocell/femtocell networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 3429–3437, Apr. 2013.

[91] Y. Huang, C. W. Tan, and B. D. Rao, "Joint beamformin and power control in coordinated multicell: Max-min duality, effective network and large system transition," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2730–2742, Jun. 2013.

[92] T. A. Le and M. R. Nakhai, "Downlink optimization with interference pricing and statistical csi," *IEEE Transactions on Communications*, vol. 61, no. 6, pp. 2339–2349, Jun. 2013.

[93] C. Lin, C. J. Lu, and W. H. Chen, "Outage-constrained coordinated beamforming with opportunistic interference cancellation," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4311–4326, Jun. 2014.

[94] W. N. S. F. Wan-Ariffin, X. Zhang, and M. R. Nakhai, "Real-time energy trading with grid in green cloud-ran," *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC 2015)*, pp. 748–752, Aug. 2015.

[95] D. Li, W. Saad, I. Guvenc, A. Mehbodniya, and F. Adachi, "Decentralized energy allocation for wireless networks with renewable energy powered base stations," *IEEE Transactions on Communications*, vol. 63, no. 6, pp. 2126–2142, Jun. 2015.

[96] N. Cordeschi, D. Amendola, M. Shojafar, and E. Baccarelli, "Performance evaluation of primary-secondary reliable resource-management in vehicular networks," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2014.

[97] H. Hindi, "A tutorial on convex optimization," *IEEE American Control Conference*, vol. 4, pp. 3252–3265, Jul. 2004.

[98] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE Journal On Selected Areas In Communications*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

[99] N. D. Sidiropoulos, T. N. Davidson, and Z. Q. Luo, "Transmit beamforming for physical layer multicasting," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.

[100] T. A. Le and M. R. Nakhai, "Coordinated beamforming using semidefinite programming," *IEEE International Conference on Communications (ICC)*, pp. 3790–3794, Jun. 2012.

[101] L. Musavian, M. R. Nakhai, M. Dohler, and A. H. Aghvami, "Effect of channel uncertainy on the mutual information of mimo fading channels," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 2798–2806, Sep. 2007.

[102] H. O. Lancaster and E. Seneta, *Chi-Square Distribution*. Wiley, 2005.

[103] A. Wiesel, Y. C. Eldar, and S. S. (Shitz), "Linear precoding via conic optimization for fixed mimo receivers," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[104] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, 2008.

[105] W. N. S. F. Wan-Ariffin, X. Zhang, and M. R. Nakhai, "Real-time power balancing in green comp network with wireless information and energy transfer," *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC 2015)*, pp. 1574–1578, Aug. 2015.

List of Publications

Journal Publications

Wan Nur Suryani Firuz Wan Ariffin, **Xinruo Zhang** and Mohammad Reza Nakhai, "Sparse Beamforming for Real-time Resource Management and Energy Trading in Green C-RAN", *IEEE Transactions on Smart Grid*, vol.8, no.4, pp.2022-2031, July 2017 [4].

Xinruo Zhang, Mohammad Reza Nakhai and Wan Nur Suryani Firuz Wan Ariffin, "A Bandit Approach to Price-Aware Energy Management in Cellular Networks", *IEEE Communications Letter*, vol.21, no.7, pp.1609-1612, July 2017 [32].

Xinruo Zhang, Mohammad Reza Nakhai and Wan Nur Suryani Firuz Wan Ariffin, "Adaptive Energy Storage Management in Green Wireless Networks", *IEEE Signal Processing Letter*, vol.24, no.7, pp.1044-1048, July 2017 [33].

Wan Nur Suryani Firuz Wan Ariffin, **Xinruo Zhang** and Mohammad Reza Nakhai, "Predictive Energy Trading in C-RAN", submitted to *IEEE Access* and under review.

Conference Publications

Wan Nur Suryani Firuz Wan Ariffin, **Xinruo Zhang** and Mohammad Reza Nakhai, "Real-time Power Balancing in Green CoMP Network with Wireless Information and Energy Transfer", *IEEE PIMRC 2015*, Aug. 2015 [105].

Wan Nur Suryani Firuz Wan Ariffin, **Xinruo Zhang** and Mohammad Reza Nakhai, "Real-time Energy Trading with Grid in Green Cloud-RAN", *IEEE PIMRC 2015*, Aug. 2015 [94].

Wan Nur Suryani Firuz Wan Ariffin, **Xinruo Zhang** and Mohammad Reza Nakhai, "Sparse Beamforming for Real-time Energy Trading in CoMP-SWIPT Networks", *IEEE ICC 2016*, May 2016 [46].

Wan Nur Suryani Firuz Wan Ariffin, **Xinruo Zhang** and Mohammad Reza Nakhai, "Combinatorial Multi-armed Bandit Algorithms for Real-time Energy Trading in Green C-RAN", *IEEE ICC 2016*, May 2016 [12].

Xinruo Zhang and Mohammad Reza Nakhai, "Robust Chance-Constrained Distributed Beamforming for Multicell Interference Networks", *IEEE ICC 2016*, May 2016 [29].

Xinruo Zhang and Mohammad Reza Nakhai, "A Distributed Algorithm for Robust Transmission in Multicell Networks with Probabilistic Constraints", *IEEE GLOBECOM 2016*, Dec. 2016 [30].

Xinruo Zhang, Mohammad Reza Nakhai and Wan Nur Suryani Firuz Wan Ariffin, "A Multi-armed Bandit Approach to Distributed Robust Beamforming in Multicell Networks", *IEEE GLOBECOM 2016*, Dec. 2016 [31].